DOCTOR OF PHILOSOPHY

Ligandability of protein-protein interactions and surfaces on Cullin RING E3 ubiquitin ligases

Cardote, Teresa Amorim de Faria

Award date:
2017

Ligandability of protein-protein interactions and surfaces within Cullin RING E3 ubiquitin ligases

Teresa Amorim de Faria Cardote

Supervisor
Professor Alessio Ciulli

Thesis submitted for the degree of
Doctor of Philosophy
School of Life Sciences
University of Dundee
May 2017

Para a minha Avó Teresa e para a minha Mãe.

To my Grandmother Teresa and to my Mother.

Acknowledgements

I would like to thank Alessio, my supervisor, for trusting me when I decided to embark in the adventure of a Ph.D. abroad. Thank you for giving me space to grow personally and scientifically. I hope I did not let you down!

I am truly thankful to Carles, Morgan, Alessio B., Michael and Kwok-Ho for their guidance and support throughout this bumpy road. You were always helpful and available, thank you.

I thank Dr. Axel Knebel and Clare Johnson (MRC-PPU) for sharing their knowledge on Cul2 with me and for their helpful attitude. I also thank my thesis committee, Prof. David Horn and Dr. Piers Hemsley for their advice and helpful discussions; all my colleagues and all the people that somehow contributed to this work. Thanks are also due to Fundação para a Ciência e Tecnologia and the Portuguese government for the Ph.D. studentship.

Being away from home was a challenge that was less hard in the company of the good friends I made over these years. Carles and Eli, thank you for your friendship and kindness. Ester, thank you for sharing with me the burden of a men-only Lab in the early days, thank you for your friendship and for smiling always. Anastasia, my partner in crime, thank you for all the support in the good and in the bad moments, and for being Bananas! It would have been a lot harder without you (and boring too!). Morgan and Jess, thank you for your friendship and for putting up with our some times creative way of speaking in English! Michy, thank you for your friendship and for fighting disorganised people alongside me. Wei-Wei, thank you for your kindness, your friendship, and your entertainment skills, of course! Gugui, thank you for your friendship (and for singing and dancing in the Lab with me!). Nicola, buddy of the Lab marathons, thank you for your company and all the grown-up talks! Tonia, Mattia, Kwok-Ho, Andrea, Xavi, Pedro, thank you for fostering a great working environment! With all of you I shared some of the best moments of the past four years, thank you for making this journey much more enjoyable.

I am really grateful to my oldest friends for being close despite the distance. Last but definitely not least, I would like to acknowledge the support of my family. Avó Teresa, Mãe, André, João, Carolina, and Renata, thank you for believing in me, trusting my choices and being always on my side. Thank you Alessio B. for your tirelessly encouragement, for calming my panic and for always looking on the bright side of life. You made it easier!

Declaration

I declare that the work described in this thesis was carried out by me under the supervision of Professor Alessio Ciulli and, unless otherwise stated, all references cited have been consulted by me. To my knowledge, the work of which this thesis is a record is original and has not been submitted for any other degree at this or other university.

Teresa Amorim de Faria Cardote

May 2017

Publications

The work described in this thesis led to the following publications.

1. Cardote, T.A.F.C., Ciulli, A., Cyclic and macrocyclic peptides as chemical tools to recognise protein surfaces and probe protein-protein interactions. *ChemMedChem* **2016**, 11 (8), 787-794.

2. Cardote, T.A.F.C., Gadd, M.S., Ciulli, A., Crystal structure of the Cul2-Rbx1-EloBC-VHL ubiquitin ligase complex. *Structure* **2017**, 25 (6), 901-911.

3. Cardote, T.A.F.C., Ciulli, A. Structure-guided design of peptides as tools to probe the protein-protein interaction between Cullin-2 and Elongin BC substrate adaptor in Cullin RING E3 ubiquitin ligases. *ChemMedChem*, manuscript accepted, DOI:10.1002/cmdc.201700359.

Summary

Cullin RING E3 ubiquitin ligases (CRLs) function in the ubiquitin proteasome system by catalysing the transfer of ubiquitin from E2 conjugating enzymes to specific substrate proteins. CRLs are large dynamic multi-subunit complexes that control the fate of many proteins in cells and, therefore, constitute attractive drug targets for the development of small-molecule tools and potential drug leads, such as inhibitors and chemical inducers of protein degradation.

This work presents the first crystal structure of the pentameric human $CRL2^{VHL}$ complex, composed of Cul2, Rbx1, Elongin B (EloB), Elongin C (EloC) and pVHL. The structure presents a closed state of full-length Cul2 and a new conformation of Rbx1, thought to be in a trajectory from inactive to active state. The thermodynamic signature of the interaction between Cul2 and pVHL-EloBC (VBC) was determined as well as mutations that contribute toward a selectivity switch for Cul2 *versus* Cul5 recognition. In addition, this work focused on an extensive approach to probe the VBC surface with peptides. A first methodology involved the structure-based design aimed at targeting the Cul2-VBC interaction. Three peptides have been shown to bind at the Cul2 binding site on EloC, however, with very weak affinities that could not be optimised, suggesting its poor ligandability. The second methodology included an unbiased screening using phage display libraries of bicyclic peptides. The screening campaign yielded peptide hits but their physicochemical properties, especially poor solubility, constituted an obstacle to the progress toward a biophysical characterisation of their binding to VBC and the development into high-affinity probes.

The findings of this work provide structural and biophysical contributions into the whole $CRL2^{VHL}$ complex assembly and functioning and provide insights and tools that could aid future targeting of this CRL.

Abbreviations

Alpha: amplified luminescent proximity homogeneous assay

ACN: acetonitrile

ASB9: ankyrin repeat and SOCS box containing 9

AU: absorption units

BLI: biolayer interferometry

CAND1: Cullin-associated NEDD8-dissociated protein 1

COP9: Constitutive photomorphogenesis 9

CPMG: Carr-Purcell-Meiboom-Gill sequence

CRL: Cullin RING Ligase

CSN: COP9 signalosome

CTD: C-terminal domain

DCM: dichloromethane

DIPEA: N,N,-diisopropylethylamine

DMF: N,N-dimethylformamide

DMSO: dimethylsulfoxide

DNA: deoxyribonucleic acid

DTT: dithiotheitrol

E. coli: *Escherichia coli*

EDT: 1,2-ethanedithiol

EloA: elongin A

EloB: elongin B

EloC: elongin C

EMA: european medicines agency

FDA: food and drug administration

Fmoc: fluorenylmethyloxycarbonyl chloride

HATU: 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate

HBTU:2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

HIF-1α: hypoxia inducible factor 1α

HECT: homologous to the E6-AP carboxyl terminus

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HOAT: 1-hydroxy-7-azabenzotriazole

HPLC: high pressure liquid chromatography

HSQC: heteronuclear single quantum coherence

IPTG: isopropyl-β-D-thiogalactopyranoside

ITC: isothermal titration calorimetry

LB: lysogeny broth

LC-MS: liquid chromatography-mass spectrometry

LOGSY: water-ligand observed via gradient spectroscopy

MBP: maltose binding protein

MES: 2-(N-morpholino)ethanesulfonic acid

MPD: 2-methyl-2,4-pentanediol

MR: molecular replacement

MS: mass spectrometry

MW: molecular weight

NaDEDTC: sodium diethyldithiocarbamate

NEDD8: neuronal precursor cell expressed developmentally down-regulated protein 8

NMM: N-methylmorpholine

NMP: N-methylpirrolidinone

NMR: nuclear magnetic resonance

NTD: N-terminal domain

OD_{600}: Optical density at 600 nm

PBS: phosphate-buffered saline

PDB: protein data bank

POI: protein of interest

PPI: protein-protein interaction

PROTAC: proteolysis targeting chimera

PTM: post-translational modification

Rbx: RING box protein

RBR: RING between RING

RFU: relative fluorescence units

RING: really interesting new gene

RMSD: root mean square deviation

RPM: rotations per minute

SDS-PAGE: sodium dodecyl sulphate polyacrylamide gel electrophoresis

SEC: size exclusion chromatography

SOCS2: suppressor of cytokine signalling 2

SSA: super streptavidin

SUMO: small ubiquitin-like modifier

TATB: 1,1',1"-(1,3,5-triazinane-1,3,5-triyl)tris(2-bromoethanone)

TB: terrific broth

TBMB: 1,5,5-tris(bromomethyl)benzene

TBMT: 2,4,6-tris(bromomethyl)-s-triazine

TBS: tris-buffered saline

TBST: TBS with 0.1% Tween-20

TCEP: tris-(2-carboxyethyl)phosphine

TEV: Tobacco etch virus protease

TFA: trifluoroacetic acid

THF: tetrahydrofuran

TIS: triisopropylsilane

TROSY: transverse-relaxation optimised spectroscopy

TSP: 3-(trimethylsilyl)propionic acid

UPS: ubiquitin proteasome system

VHL: von Hippel–Lindau protein

WT: wild type

Table of contents

Acknowledgements		V
Declaration		VII
Publications		IX
Summary		XI
Abbreviations		XIII
Table of contents		XVII

CHAPTER 1: INTRODUCTION

1.1	Motivation	25
1.2	The ubiquitin proteasome system	27
1.2.1	Cullin RING E3 ubiquitin ligases	29
1.2.1.1	The VHL ligase: $CRL2^{VHL}$	31
1.3	Targeting protein-protein interactions	35
1.3.1	Peptides as tools to explore PPIs and surfaces	37
1.4	Biophysical techniques	39
1.4.1	Protein X-ray crystallography	39
1.4.2	Differential scanning fluorimetry	40
1.4.3	Isothermal titration calorimetry	41
1.4.4	Nuclear magnetic resonance	41
1.4.5	Biolayer interferometry	43
1.4.6	AlphaLISA	44
1.5	Prior art	45
1.6	Aims	47

CHAPTER 2: STRUCTURAL AND BIOPHYSICAL INSIGHTS INTO THE CRL2^{VHL} MULTI-SUBUNIT PROTEIN COMPLEX

2.1	Introduction	51
2.2	Results	53
2.2.1	Protein expression and purification	53
2.2.2	Crystal structure of the CRL2^{VHL}	55
2.2.2.1	Cullin-2 structure highlights its inherent flexibility	62
2.2.2.2	Rbx1 presents a new orientation	64
2.2.2.3	The interface between Cul2 and VBC	66
2.2.3	Neddylation of the CRL2^{VHL} complex	68
2.2.4	Biophysical characterisation of the Cul2–VBC interaction	69
2.2.5	Critical residues for the Cul2–VBC interaction	72
2.2.5.1	Site-directed mutagenesis, protein expression and purification	74
2.2.5.2	AlphaLISA experiments	74
2.2.5.3	ITC experiments	75
2.2.6	Swapping residues and selectivity for Cul2 *versus* Cul5	76
2.2.6.1	Site-directed mutagenesis, protein expression and purification	78
2.2.6.2	AlphaLISA and ITC experiments	78
2.3	Discussion	82

CHAPTER 3: PROBING THE CRL2^{VHL} COMPLEX

3.1	Introduction	89
3.2	Results	90
3.2.1	Design, expression and optimisation efforts to obtain soluble and monomeric Cul2$_{NTD}$ constructs	90
3.2.1.1	Expression and purification of Cullin-2$_{1-163}$-MBP-6xHis	91
3.2.1.2	Expression and purification of Cullin-2$_{1-384}$-MBP-6xHis	96
3.2.2	3- to 11-mer Cul2 peptides	103

3.2.2.1	Peptide synthesis and purification	103
3.2.2.2	Biolayer interferometry	103
3.2.3	8-mer peptide derivatives	106
3.2.3.1	Alanine scanning	106
3.2.3.2	Leucine replacement	108
3.2.3.2.1	Characterisation of peptide **J** binding to VBC	111
3.2.4	Cyclic pentapeptides	114
3.2.4.1	Peptide synthesis and purification	115
3.2.4.2	Biophysical characterisation	118
3.2.5	Peptides targeting the charge complementarity interface	118
3.2.5.1	Biophysical characterisation	118
3.2.6	Clamp peptides	120
3.2.6.1	Biophysical characterisation	122
3.3	Discussion	123

CHAPTER 4: BICYCLIC PEPTIDES AND STUDIES OF THEIR BINDING TOWARDS THE VBC COMPLEX

4.1	Introduction	131
4.2	Results	134
4.2.1	Phage display screening	134
4.2.1.1	Screening of the VL1 library	136
4.2.1.2	Screening of the VL2 library	137
4.2.1.3	Screening of the VL3 library	138
4.2.1.4	Screening of VL1, VL2 and VL3 together	139
4.2.2	Synthesis of the bicyclic peptides identified from the phage display screening	141
4.2.3	Initial screening of the bicyclic peptides	143
4.2.3.1	Differential Scanning Fluorimetry	143

4.2.3.2	Biolayer interferometry	146
4.2.3.2.1	Single-point concentration assay	146
4.2.3.2.2	Concentration-dependent response assays	146
4.2.4	Large-scale synthesis and purification of the hit peptides	149
4.2.5	Validation of binders	151
4.3	Discussion	155

CHAPTER 5: CONCLUDING REMARKS

5.1	Concluding remarks	161

CHAPTER 6: EXPERIMENTAL

6.1	Molecular biology methods	167
6.1.1	Plasmids	167
6.1.2	Cloning of the truncated Cul2 NTD constructs	169
6.1.3	Site-directed mutagenesis	170
6.1.4	Transformation	171
6.1.5	BL21 pGro7 competent cells generation	171
6.1.6	Preparation of DH10Bac competent cells	172
6.1.7	Bacmid generation	172
6.1.8	Insect cells tissue culture	173
6.1.9	Bacmid transfection into *Sf21* cells	173
6.1.10	Baculovirus amplification	174
6.1.11	Protein expression in bacteria	174
6.1.12	Protein expression in insect cells	178
6.1.13	Isotopically-labelled VBC expression	183
6.1.14	Biotinylation of proteins	185
6.1.15	DNA and protein concentrations	185

6.1.16	Immunoblotting	185
6.1.17	Neddylation assay	186
6.1.18	Preparation of ampicillin-modified sepharose resin	186
6.2	Crystallography	187
6.2.1	$CRL2^{VHL}$ complex crystallisation	187
6.2.1.1	Data collection and structure solving	188
6.2.2	VBC crystallisation and soaking	188
6.3	Phage display	189
6.3.1	Library amplification	189
6.3.2	Peptide cyclisation	189
6.3.3	Biopanning	190
6.4	Peptide synthesis and characterisation	191
6.4.1	Cullin-2 peptides, 15-mer peptide and clamp peptides	192
6.4.2	Cyclic pentapeptides	207
6.4.3	Bicyclic peptides	208
6.5	Biophysical techniques	212
6.5.1	Isothermal titration calorimetry	212
6.5.2	Differential scanning fluorimetry	212
6.5.3	Biolayer interferometry	213
6.5.4	NMR spectroscopy	213
6.5.5	AlphaLISA	215

CHAPTER 7: REFERENCES

APPENDIX I: Supplementary data 243

APPENDIX II: Publications 269

CHAPTER 1

INTRODUCTION

1.1 Motivation

In drug discovery, the ligandability of a target is a very critical property. Ligandability is defined as the propensity of a target to bind small molecules,[1] and it constitutes a requirement for druggability *i.e.* the modulation of a target by drug-like small molecules as a mechanism to elicit a therapeutic benefit.[2] In order to be druggable, a target must be ligandable, however, being ligandable, despite increasing the chances, does not always guarantee target druggability. Nowadays, different avenues are pursued by the research community to explore the ligandability of potential targets, before embarking on costly and lengthy drug screening campaigns. Fragment screening approaches to identify chemical starting points for optimisation into more potent ligands, typically allied with structural and biophysical studies of a particular target, represent a prominent and powerful way to probe target ligandability.

Traditionally, proteins are targeted at well-defined ligand binding sites or catalytic grooves and these conventional targets are mainly enzymes and protein receptors, such as kinases and G-protein-coupled receptors (GPCRs). These key families of proteins constitute only a small portion of the whole proteome, highlighting that a great amount of it remains untargeted. In fact, only 10-14% of the human genome is predicted to be druggable.[2] To expand the druggable portion of the genome, the focus must shift to unconventional targets. These comprise, for example, proteins that endogenously bind peptides or other proteins as ligands, or that are involved in protein-protein interactions (PPIs), and that are usually deemed undruggable. This is due to their lack of small-molecule binding sites, which are often replaced by large and flat surfaces.

In fact, most of the proteins do not function alone, indeed, their function depends on the interaction with other proteins, and PPIs are biologically relevant as regulators of most biochemical pathways. The human interactome (set of PPIs that occur in a cell), is estimated to comprise around 100,000 protein interactions,[3] and each of this PPI provides an opportunity for therapeutic intervention. Therefore, despite challenging, PPIs and surfaces provide exciting new opportunities to be explored as potential targets.[4]

The challenges in assessing the ligandability of PPIs and protein surfaces represent usually a high risk-benefit balance. A protein surface might be ligandable but the ligand site might not constitute a PPI. At the same time, by identifying a novel ligand site one can contribute to unveiling unknown native partners. A complete model system to assess ligandability of a target and explore new and known binding sites in a combined fashion is a multi-subunit protein complex. On one hand, it is a system that relies on intrinsic PPIs for function (known binding sites) and, on the other hand, it is likely that each of the subunits might be ligandable on its own (new binding sites). Apart from being model systems, studying multi-subunit complexes is essential to understand key cell physiology features due to their frequent involvement in critical cellular processes, including cell signalling, gene expression, metabolism and cell cycle regulation.[5] However, several aspects must be contemplated. Probing a large complex can be a demanding task in terms of biochemical and biophysical techniques available and their limitations, as, for example, most of the traditional methods are optimal for proteins under 50 kDa.[6] Equally, the independent study of the individual subunits of the complex can prove difficult, as the proteins might not be stable on their own, since they naturally function bound to other proteins and tend to have large hydrophobic surfaces. Overall,

only the consolidation of insights gathered from different techniques and methodologies can assist toward the final goal of understanding the ligandability of a multi-subunit protein target.

The motivation behind this work contemplates a multi-disciplinary approach to address the challenges presented above by investigating the ligandability of protein-protein interactions and surfaces within Cullin RING E3 ubiquitin ligases.

1.2 The ubiquitin proteasome system

In living organisms, complex signalling mechanisms and networks continuously regulate cells. The rates of protein degradation are an important feature of protein homeostasis that impact upon cell regulation and half-lives of proteins within cells vary greatly according to their function. For example, transcription factors have short half-life times as the rapid turnover of these proteins is needed in order to allow their levels to rapidly change in response to external stimuli. Additionally, faulty proteins are, in healthy systems, detected and rapidly degraded to limit the consequences of these defects.[7] Protein degradation in eukaryotic cells is mediated by two major pathways: the ubiquitin proteasome system and lysosomal proteolysis.[8]

The discovery of the ubiquitin proteasome system (UPS) was a landmark in the understanding of protein regulation in cells and its importance has been acknowledged with the award of the Nobel Prize in Chemistry to Avram Hershko, Aaron Ciechanover and Irwin Rose in 2004.[9,10] Functioning of the UPS is dependent on the sequential action of three key players: the E1 activating enzyme, the E2 conjugating enzyme and the E3 ligase enzyme (Figure 1.1). The role of the E1 enzyme consists of the activation of ubiquitin, a

highly abundant 76 amino acid protein, by forming a thioester bond with it through a cysteine residue. Next, activated ubiquitin is transferred to a cysteine residue in the active site of the E2 conjugating enzyme. Finally, the E3 ligase enzyme catalyses the transfer of ubiquitin from the loaded E2 enzyme to the substrate protein. Ubiquitin is linked to the substrate through an isopeptide bond between its C-terminal and a ε-amino group of a specific lysine of the target protein.[10]

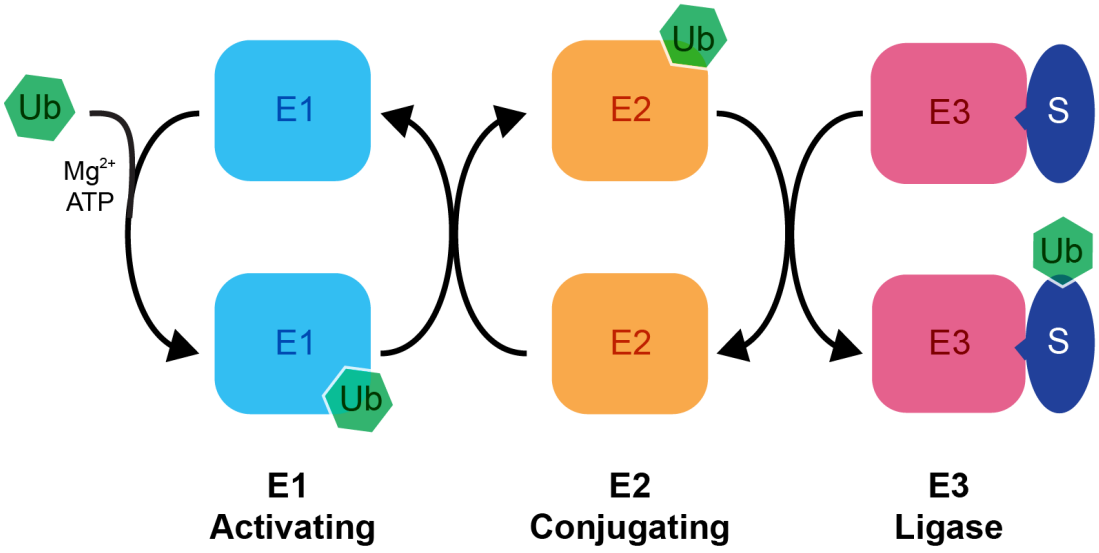


Figure 1.1 – The ubiquitin proteasome system. The ubiquitin proteasome system depends on the concerted action of three enzymes: the E1 activating enzyme, the E2 conjugation enzyme and the E3 ligase enzyme. Ubiquitin (Ub) is activated by the E1 enzyme and then transferred to the E2 enzyme. The E3 ligase catalyses the transfer of ubiquitin from the E2 enzyme to the substrate (S).

The human genome encodes for two E1 enzymes, at least 38 E2 enzymes and over 600 E3 enzymes.[11,12] The E3 ligases are responsible for recruiting the substrates to be ubiquitinated and confer specificity to the system. Proteins can be monoubiquitinated, multi-monoubiquitinated or polyubiquitinated and the ubiquitin chain can be linear, branched or mixed (Figure 1.2). The length and topology of the ubiquitin chain attached to the protein substrate greatly determines the cellular response triggered.[13] The outcome of this tagging event

can vary from activation of the substrate for proteasomal or lysosomal degradation, to direct modulation of protein interactions, activity or protein localisation.[13]

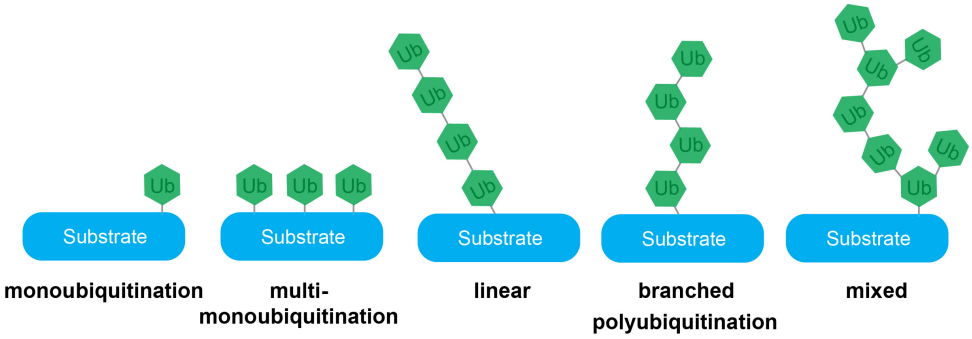


Figure 1.2 – Length and topology of the ubiquitin chain. The number of ubiquitin molecules as well as the way they are attached to the substrate determine the cellular response to this tagging event. Substrates can be monoubiquitinated once or multiple times and can be polyubiquitinated in a linear, branched or mixed fashion.

1.2.1 Cullin RING E3 ubiquitin ligases

Cullin RING E3 ubiquitin ligases (CRLs) are the focus of this work. They constitute one of the three main families of E3 ubiquitin ligases currently known. HECT and RBR E3 ligases are the other two families and despite all carrying out the last step of ubiquitination, the different classes vary in terms of structure and mechanism of ubiquitin transfer.[14] Whilst CRLs catalyse the direct transfer of ubiquitin from the E2 conjugating enzyme to the substrate, HECT and RBR ligases contain themselves a catalytic cysteine that forms a thioester intermediate with ubiquitin prior to transfer to the substrate.[15] CRLs, of which over 230 are estimated in human cells, are responsible for approximately 20% of the ubiquitin-dependent protein turnover in cells,[16] being implicated in a number of cellular processes across different organisms (Figure 1.3).[17]

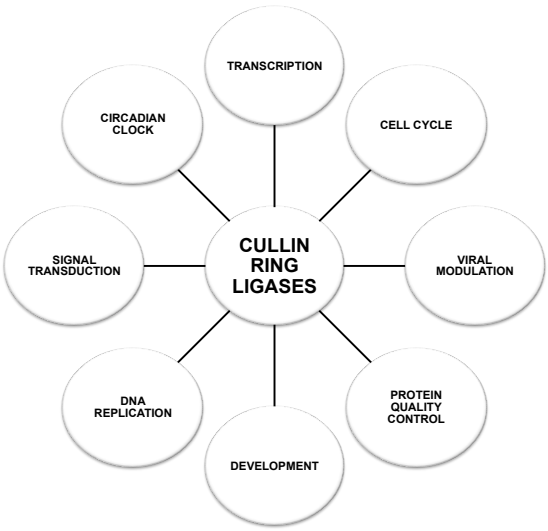


Figure 1.3 – Roles of CRLs in diverse cellular processes. CRLs are implicated in a variety of cellular processes in different organisms.

Structurally, CRLs are organised around a central scaffold subunit, Cullin, that brings together the substrate to be ubiquitinated and the E2 enzyme loaded with ubiquitin. The substrate is recruited through a receptor subunit, whereas the E2 enzyme is recruited through the Rbx RING domain. There are seven characterised Cullin proteins, Cul1, Cul2, Cul3, Cul4A, Cul4B, Cul5 and Cul7 (Cul9 is the eighth member of the family, although little is still known about this protein) and each of them can assemble with a wide variety of receptors, originating E3 ligases that, despite sharing the scaffolding core, recruit different substrates.[18] This combinatorial arrangement is a key aspect for these enzymes, accounting for their modularity and diversity of substrates that are ubiquitinated. Diversity and specificity are further increased by post-translational modifications (PTMs) that are often a requirement for substrate recognition by the CRLs, and in some cases other protein or small-molecule factors.[16] Apart from the regulation of CRLs activity by PTMs, which dictate when ubiquitination occurs, other mechanisms are involved in the regulation of the catalytic core. For instance, the attachment of NEDD8, a ubiquitin-like protein, to a specific

lysine residue in Cullin (neddylation) is an important step in the activation of the enzymatic activity.[19] Cullins undergo cycles of neddylation and de-neddylation as part of their function and regulation. Neddylation is accomplished by specific E2 enzymes that bind the RING domain of CRLs and transfer NEDD8 to Cullin, whereas de-neddylation is a reaction catalysed by the COP9 signalosome (CSN), also a large multi-subunit complex (Figure 1.4B).[20] Once de-neddylated, CRLs are sequestered by binding of CAND1 (Cullin-associated NEDD8-dissociated protein 1) that keeps them in an inactive state (Figure 1.4B).[21,22] In Cul1, for example, CAND1 competes with the substrate receptor, Skp1, for binding and it binds only to unneddylated Cullin.[23]

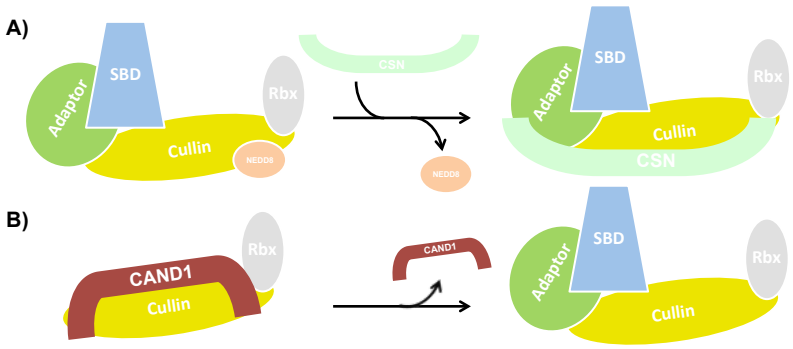


Figure 1.4 – Mechanisms of regulation of CRLs. A) De-neddylation by the COP9 signalosome (CSN). **B)** Sequestering by CAND1 keeps Cullin in an inactive state.

The important roles of CRLs in several biological processes and human diseases have rapidly emerged, in particular in cancer, as the genes encoding for E3 ligase subunits and their native substrates are often found as oncogenes or tumour suppressors.[24] These, amongst other motivations, justify the growing interest in this class of enzymes for research and therapy.[25]

1.2.1.1 The VHL ligase: CRL2^{VHL}

The CRL complex investigated in this work is the VHL ligase, CRL2^{VHL}. The central scaffold of the VHL ligase is Cul2, which recruits at the N-terminal

domain the von Hippel-Lindau protein (pVHL) as substrate receptor, through an adaptor subunit constituted by Elongin B (EloB) and Elongin C (EloC) and at the C-terminal domain the RING box protein, Rbx1 (Figure 1.5).

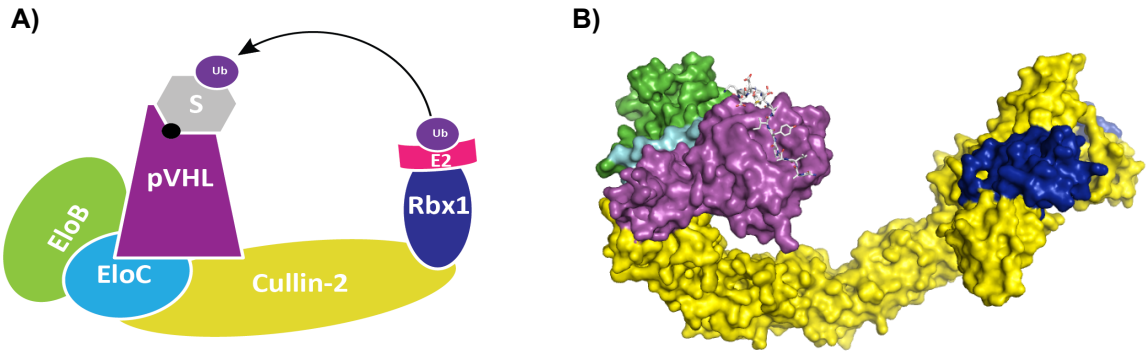


Figure 1.5 – The CRL2^{VHL} complex. A) Diagram model of CRL2^{VHL}. **B)** Structural model of CRL2^{VHL}. The pentameric complex is organised around the central scaffold, Cul2, which brings together the substrate and the E2 conjugation enzyme loaded with ubiquitin. The substrate (S), shown as HIF-1α peptide on **B**), is recruited by pVHL, which binds to the N-terminal domain of Cul2 through the adaptor subunit composed of EloB and EloC. The E2 enzyme is recruited to the complex through the RING domain of Rbx1. Model obtained using PBD 4AJY[52] for VBC in complex with HIF-1α peptide.

pVHL is a tumour suppressor protein. Mutation of both alleles of the *VHL* gene result in the von Hippel-Lindau syndrome, a relatively rare hereditary cancer syndrome characterised by a variety of tumours, being the most frequent hemangioblastomas of the central nervous system and retina, renal cells carcinoma and pheochromocytoma.[26] While its function as the substrate receptor of CRL2^{VHL} is the best characterised, studies suggest that pVHL has also E3 ligase independent functions, where it interacts and regulates other proteins such as fibronectin or collagen, for example, but does not promote their polyubiquitination and consequent degradation.[27]

The dimer composed by EloB and EloC (EloBC) acts as an adaptor between the substrate receptor and the scaffold in Cul2 and Cul5 containing CRLs. Apart from its function in E3 ubiquitin ligases, EloBC forms a trimeric

complex with Elongin A (EloA), functioning as positive regulator of the RNA polymerase II.[28] pVHL binds to EloBC through a conserved motif, referred to as BC-box. VHL disease and sporadic clear cell renal carcinoma are, in a large number of cases, promoted by deletion or mutation of the BC-box, precluding pVHL to bind EloBC.[29]

Cul2 is the scaffolding core of the VHL ligase and has a vital role in cells, as supported by genetic evidence. It is anticipated that full knockout of Cul2 may cause embryonic lethality, given that it is responsible for the ubiquitination and degradation of proteins with important functions in cell cycle and progression – and indeed there are no published studies with mouse knockouts of Cul2.[30] Mouse knockouts of Cul1, Cul3, and Cul4B resulted in embryonic lethality, suggesting indispensable roles of these proteins during early development. Cul4A knockout resulted in skin carcinogenesis, proliferation defects and male infertility, whereas the knockout of Cul7 resulted in birth lethality. On the other hand, Cul9 knockout resulted in a viable embryo, suggesting this protein is not needed for embryonic development.[30]

The last member of the complex, Rbx1, is a RING box protein essential for the catalytic activity of $CRL2^{VHL}$ and many other CRLs, being implicated in both neddylation and ubiquitination reactions. Rbx1 is essential during mouse embryonic development[30] and is overexpressed in multiple human cancer tissues.[31] Its function in the complex is to recruit the E2 conjugating enzyme loaded with NEDD8 or ubiquitin. Rbx1 has been also shown to function as NEDD8 ligase for Rub1 (the yeast homolog of NEDD8).[32]

At least six proteins have been identified as substrates of the $CRL2^{VHL}$ complex.[33] Sprouty2 (Spry2) regulates cell migration and proliferation in response to growth factors and is recognised by pVHL for degradation.[34]

Another known substrate is Rbp1, the largest subunit of the RNA polymerase II. Rbp1 is hydroxylated under oxidative stress conditions, recruited by pVHL and polyubiquitinated by the $CRL2^{VHL}$ complex. It is thought that after being ubiquitinated, the fate of Rbp1 is determined by the cellular context.[35,36] Similarly, another subunit of the RNA polymerase II, Rpb7, has been reported to be degraded after being polyubiquitinated by the VHL ligase.[37] The atypical protein kinase C (PKC) is another identified substrate recruited by pVHL, which is equally polyubiquitinated and degraded by the proteasome. In addition, the epidermal growth factor receptor (EGFR), has also been shown to be targeted by the $CRL2^{VHL}$ ligase for degradation.[38] The mediated degradation of the hypoxia inducible factor 1 alpha subunit (HIF-1α) is, however, the most widely studied and best characterised function of the $CRL2^{VHL}$ ligase. Oxygen concentration tightly regulates HIF transcriptional activity. Under normal levels of oxygen (normoxia) HIF-1α is efficiently hydroxylated by oxygen-dependent prolyl hydroxylase domain (PHD) enzymes at specific proline residues (Pro402 and Pro564) within the so-called oxygen destruction domain (ODD) of HIF-1α, which functions as a degron.[39] The hydroxyproline PTM is recognised via the β domain of pVHL and HIF-1α is recruited to the $CRL2^{VHL}$ complex to be ubiquitinated and targeted for degradation by the proteasome.[40] Upon decrease of the oxygen levels (hypoxia), PHD activity is inhibited and HIF-1α is no longer hydroxylated, escapes E3 recognition and accumulates in the cell, triggering a transcriptional response to hypoxia.[41] Stabilisation of HIF-1α in cells has proven beneficial in diseases such as chronic anaemia and acute ischemic disorders and for this reason the $CRL2^{VHL}$ complex is considered an attractive therapeutic target.[42] Work by other members of the Ciulli laboratory has recently generated potent pVHL inhibitors using structure-guided drug design.[43] Compound VH298

was characterised in cells as a highly selective inhibitor active on-target against pVHL, and elected as a chemical probe of the hypoxia signalling pathway.[44] Using these pVHL-targeting ligands as a starting point, bivalent molecules have been designed, which recruit proteins into proximity of the $CRL2^{VHL}$ to deplete protein levels inside cells and *in vivo* (Figure 1.6).[45] Targeting $CRL2^{VHL}$ to induce protein degradation provides an attractive chemical biology approach for target validation[46,47] and a new therapeutic modality in drug discovery.[48]

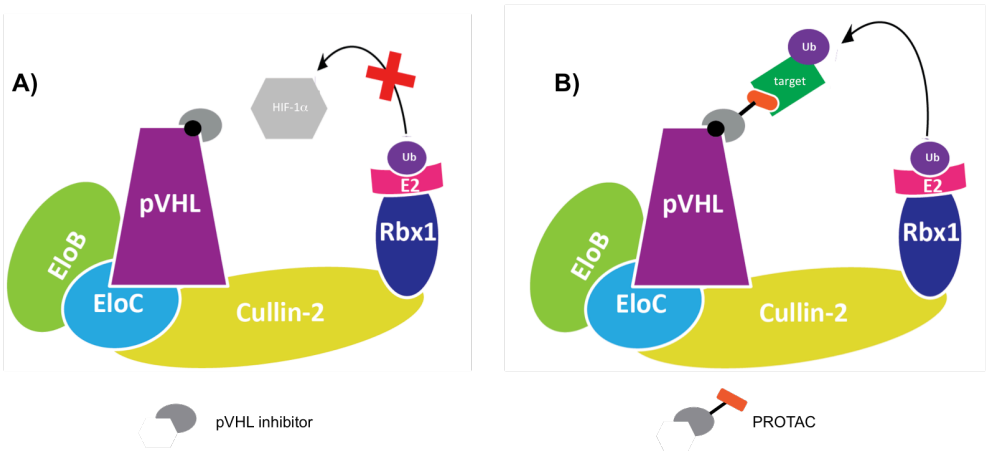


Figure 1.6 – Targeting the $CRL2^{VHL}$ complex. A) Inhibitors of the pVHL-HIF-1α interactions. **B)** Bivalent molecules (proteolysis targeting chimeric molecules, PROTACs) that recruit specific proteins to the $CRL2^{VHL}$ complex to be ubiquitinated and targeted for degradation.

1.3 Targeting protein-protein interactions

Protein-protein interactions are involved in the control and regulation of cellular processes that range from enzyme catalysis to cell signalling and development and to protein homeostasis. PPIs have a strong influence in the cellular localisation and abundance of the proteins involved, and often determine the fidelity and specificity of their function.[49] The highly organised and responsive protein networks that regulate most biological processes further underscore the importance of PPIs, providing many opportunities for therapeutic intervention. In fact, each PPI in such networks may play a role in physiological and pathophysiological states, and could thus represent a

prospective drug target with potential therapeutic relevance.[50] In this context, the development of chemical tools or probes that can help the understanding of mechanisms and biological roles of specific PPIs is of extreme relevance.

Despite the great research and therapeutic opportunities of targeting PPIs, the physicochemical nature of these interfaces makes it a challenge to develop small molecules that bind to these sites and potentially disrupt these contacts. PPIs can be classified according to their binding epitope as primary, secondary and tertiary.[51] In primary PPIs, a linear peptide binds to a protein surface, and usually the binding event is driven by recognition of a PTM; in secondary PPIs, a peptide with secondary structure binds to a protein surface; and finally, tertiary PPIs involve multiple contact points and require a tertiary structure (Figure 1.7).

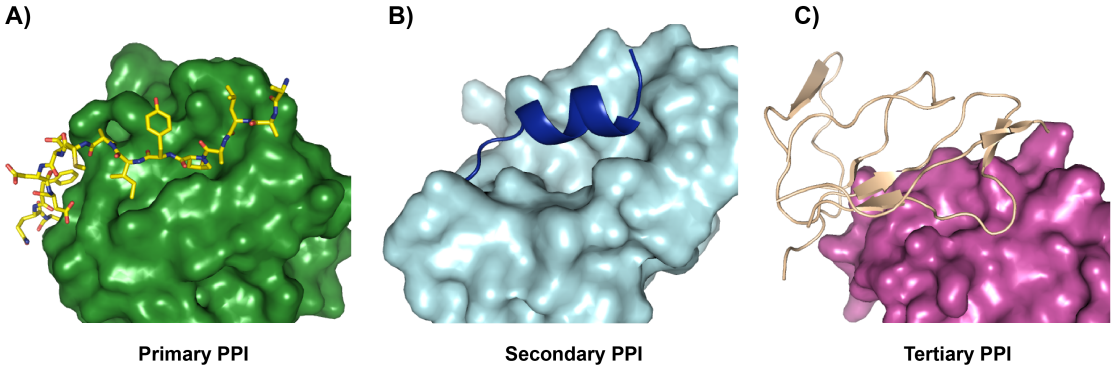


Figure 1.7 – Classes of protein-protein interactions. A) Primary protein-protein interaction between HIF-1α peptide in yellow and pVHL in green (PDB: 4AJY).[52] **B)** Secondary protein-protein interaction between MDM2 in cyan and p53 in blue (PDB: 1YCR).[53] **C)** Tertiary protein-protein interaction between IL-2 in salmon and IL-2Rα in magenta (PDB: 1Z92).[54]

Contrary to the more traditional binding sites defined at the active sites of receptors or enzymes, which tend to be buried from solvent and contain well-defined and deep pockets, PPIs usually involve flat surfaces of extended area, or in any case smaller pockets to merely accommodate protein side

chains.[55] Therefore, the identification and development of binding ligands to protein surfaces, whether direct or allosteric modulators of PPIs, remains a difficult and unsolved problem. Fortunately, much progress has been made in recent years to tackle this issue. In particular, it is becoming increasingly clear that the development of drug-like PPI inhibitors, and small-molecule ligands to protein surfaces, greatly benefits from the availability of a peptidic ligand to that binding site, either from the natural interacting partner or from synthetic sources.

1.3.1 Peptides as tools to explore PPIs and surfaces

Peptides are good candidates to accomplish effective targeting of PPIs due to their ability to closely mimic many structural features of protein interfaces, as described above. Some remarkably successful examples of using peptides from the natural protein partner as a starting point to drug design include well-characterised systems such as MDM2-p53[56,57] or pVHL-HIF1α.[43,52,58] Such peptidic ligands can inform on the nature and details of key interactions essential to achieve affinity and selectivity at the targeted binding site. They also provide important displacement tools to ensure specificity of interaction of compound series in ligand development processes. Furthermore, peptides constitute an interesting alternative in their own right as PPI modulators, presenting a number of advantages over non-peptidic small-molecules: biocompatibility, and low toxicity to the organism; chemical flexibility, such as the ability to adapt to large and often flexible surfaces; modularity, thus enlarging the structural diversity, enhancing selectivity and leading to high potency.[59] Notwithstanding, peptides present low stability in plasma, cleavage by proteolytic cellular enzymes and *in vivo* clearance by hepatic and renal

activities, which represent major drawbacks for their use as drugs, especially when compared to proteins, antibodies or drug-like compounds.[60] These limitations are largely associated with the flexible conformation adopted by short peptides in solution.

One way of stabilising the secondary structure of short peptides is by using a stapling approach.[61] Stapled peptides involve cross-linking at different peptide positions bridging regions proximal in space, and are often designed based on the crystal structure of a certain PPI.[62] Conformational stabilisation in particular of α-helical peptides by hydrocarbon stapling has proven successful in different systems with the resulting molecules acting as antiviral, antimicrobial and anticancer, modulators of signalling cascades and modulators of transcription and translation processes.[61]

Another way of constraining the conformation of peptides and hence develop them into molecules with more adequate structural features and physicochemical properties for protein targeting is by covalent cyclisation. Cyclic peptides present favourable binding properties by reducing the entropic penalty upon binding. In addition, their large surface area provides opportunities to achieve high target binding affinity and selectivity.[63] The constrained structure of cyclic peptides provides also increased proteolytic stability *in vivo* when compared to their linear counterparts.[64] Cyclic peptides can present different formats and arrangements, therefore, the term is general and it encompasses several classes of peptidic molecules.[65] Bicylic peptides, for example, are one of the cyclic peptides subclasses and can be obtained by the covalent attachment of two flexible peptidic loops onto a central rigid scaffold.[66] Generally, these peptides present improvements in binding affinities as well as

greater resistance to intracellular degradation and metabolic activities compared to monocyclic peptides. [67]

The potential of cyclic peptides is underpinned by the number of derived drugs being approved or in clinical studies by the Food and Drug Administration (FDA) and the European Medicines Agency (EMA), as is the case of oritavancin, an antibiotic for bacterial infections, or romidepsin, an anti-cancer agent, for example.[63]

1.4 Biophysical techniques

The interrogation and characterisation of PPIs and protein surfaces is a challenging aim that relies deeply on biochemical, biophysical and structural biology techniques.[6] A brief introduction to the biophysical techniques used throughout this thesis to characterise protein-protein or ligand-protein interactions is given below. In-depth coverage of all techniques available to study PPIs and binding of small molecules is beyond the scope of this thesis, and the reader is referred to recent reviews on this topic.[68,69]

1.4.1 Protein X-ray crystallography

X-ray crystallography is an essential technique for structure determination of protein and protein-ligand complexes that relies on the scattering of X-rays by electrons. It requires the object of analysis to be crystallisable and to produce well-diffracting crystals.[69] The result of the exposure of a crystal to X-rays is a diffraction pattern that holds information about the tri-dimensional structure of the protein or protein-ligand complex. Subsequent data processing allows the reconstruction of an electron density map that can be interpreted in terms of individual atoms and molecules. From

this interpretation, an atomic or molecular model can be generated and refined to reach the best agreement between the observed reflection amplitudes (F_{obs}) and those calculated from the model (F_{calc}). This agreement can be evaluated by a number of indicators, being the one most widely used the crystallographic R-factor.[70]

A critical aspect of X-ray crystallography that determines the extent of information retrieved from an experiment is the resolution at which the crystal diffracts X-rays. Typically, large molecules or large multi-subunit complexes diffract X-rays to resolutions greater than 3.5 Å, which constitutes a hindrance to the field, as solving and refining structures at low resolution is still a challenging task.[71] However, the increasing number of low-resolution structures deposited in the Protein Data Bank (PDB) over the years proves the experimental and computational advances for structure determination that help tackling the problems associated with low-resolution data.[72] Therefore, biological information that can only be obtained from low-resolution data is now valued and the threshold for acceptability has changed with a number of significant structures solved at low resolution, as is the case of the human DNA dependent protein kinase catalytic subunit (solved at 4.3 Å)[73] or the structure of the human antibodies 100 and 114 in complex with the Ebola virus fusion glycoprotein (solved at 6.7 Å),[74] for example.

1.4.2 Differential scanning fluorimetry

A differential scanning fluorimetry (DSF) experiment determines the thermal stability of a protein under certain conditions. The protein solution is mixed with an environmentally-sensitive fluorescent dye, SyproOrange™, and the temperature of the sample is increased gradually, typically from 25 to 95 °C.

This is monitored by continuous fluorescence readings. Once the protein is denatured, it unfolds and exposes hydrophobic surfaces. The dye binds non-specifically to these hydrophobic patches, leading to an increase in the measured fluorescence (relative fluorescence units, RFU). The plot of RFU *versus* temperature originates a typically sigmoidal curve. The melting temperature (T_M) is defined as the temperature at which half of the protein is denatured and can be calculated as the inflection point of this curve – or the minimum of the derivative function (-dRFU/dT). In the presence of a stabiliser, the T_M of a protein is expected to increase and vice-versa, a destabiliser decreases it.[75]

1.4.3 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) is a label-free technique that allows the thermodynamic characterisation of bimolecular interactions. The principle relies on the direct measurement of the heat absorbed or released upon the mixing of two solutions, each containing one species of interest. If the two species bind to each other, the measured heat signal will include a contribution from the binding event. By measuring the heat change throughout a titration, ITC enables direct and accurate determination of the binding constant (K_a), the stoichiometry of the reaction (N value), and the change in enthalpy (ΔH). The change in entropy (ΔS) can then be calculated from $\Delta G = \Delta H - \mathrm{T}\Delta S$. Altogether, these parameters provide a full thermodynamic profile for a given interaction.[76]

1.4.4 Nuclear magnetic resonance

Nuclear magnetic resonance (NMR) is a powerful technique that enables the study of ligand-protein interactions over a wide range of affinities. In this

technique, changes in the chemical shift, nuclear overhauser effects (NOEs), diffusion constants or relaxation times are measured and used to inform about the nature of a certain bimolecular interaction. The experiments can be ligand-observed or protein-observed, depending on which resonances are being monitored. In both cases, the resonance peaks are analysed and expected to change as a result of a binding event (Figure 1.8).[77]

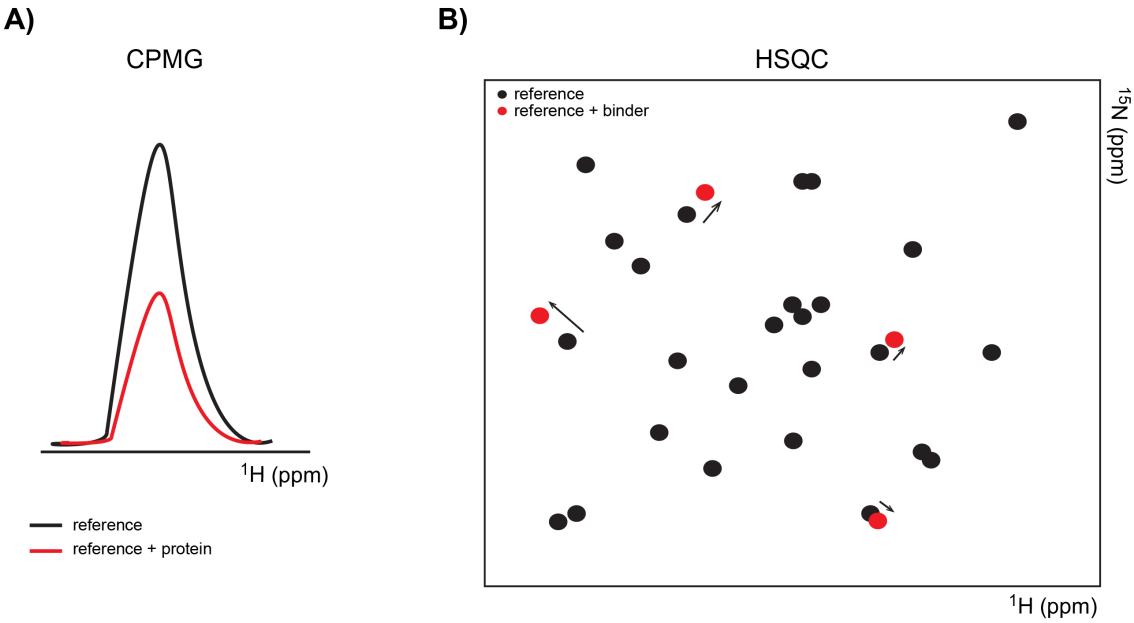


Figure 1.8 – Nuclear magnetic resonance experiments. A) Ligand-observed NMR. Example of a CPMG experiment - in the presence of a protein that binds the ligand, the signal of the ligand is expected to decrease in intensity. **B)** Protein-observed NMR. HSQC spectra, in which peaks corresponding to N-H groups of residues affected by the binding of a ligand are shifted compared to a reference spectrum.

Despite the existence of several ligand-observed experiments, only one was used in this work: relaxation-edited Carr-Purcell-Meiboom-Gill (CPMG). CPMG relies on the different re-orientational frequencies of small-molecules and macromolecules in solution (which result in different T2 transverse relaxation times). Upon binding of a protein to a small molecule, the relaxation time of the latter decreases. This is translated in a reduction of the

small-molecule signal in a spectrum of sample containing protein-ligand complex relative to a spectrum of the small molecule alone (Figure 1.8A).[78]

Protein-observed NMR relies on changes in the chemical environment of interacting protein residues upon a binding event, that translate into a change in the chemical shift of the residues involved in the interaction. In order to monitor the changes in the chemical shifts of the target protein, a two-dimensional ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) spectrum is recorded in the absence (reference) and in the presence of a ligand. The residues affected by the binding of the ligand to the protein will present a different chemical shift than in the reference spectrum (Figure 1.8B). These changes, known as chemical shifts perturbation (CSP) allow the mapping of a binding site on a protein, as long as the peaks are assigned to the corresponding residue.[79]

1.4.5 Biolayer interferometry

Biolayer interferometry (BLI) is a label-free technique that relies on the principle of optical interferometry. BLI measures molecular interactions by analysing the interference pattern of white light reflected from two surfaces – a layer of immobilised protein on the biosensor tip and an internal reference layer. Changes in the number of molecules bound to the biosensor tip cause a shift in the interference pattern that can be measured in real-time. Only molecules binding or dissociating from the biosensor are able to shift the interference pattern and thus generate a response – unbound molecules or changes in the refractive index of the surrounding medium do not affect the interference pattern. From measuring dose-response curves it is possible to extrapolate a binding affinity (K_d) for a particular interaction. In addition, BLI experiments can

provide useful kinetic parameters such as on and off-rate constants (k_{on} and k_{off}).[80]

1.4.6 AlphaLISA

In the alphaLISA experiments (amplified luminescent proximity homogeneous assay) two interaction partners are individually attached to donor and acceptor beads. Upon irradiation with light at 680 nm, the donor beads, containing a photosensitiser (phtalocyanine), produce singlet oxygen excited state ($^1O_2$) from environment oxygen (O_2). Singlet oxygen diffuses up to approximately 200 nm during its half-life time of approximately 4 μs. When the two partners are brought together in space by binding to each other, the beads come to proximity and the singlet oxygen reacts with a thioxene moiety in the acceptor bead, resulting in emission of light at around 615 nm (Figure 1.9A).[81] AlphaLISA can be used for measuring direct interactions between partner molecules, however, the multiplicity of the bead derivatisation generates multiple binding sites close to each other and, as a result, multivalent interactions that artificially increase the signal but also the affinity between binding partners. This makes the technology unsuitable to measure accurately direct binding K_d values between partners attached to the beads (Figure 1.9A), however, it makes it ideal to develop screening assays as competitive experiments (Figure 1.9B). In this format, a competitor ligand is titrated to measure its ability of disrupting the interaction between the two bead-bound partners. Upon effective displacement, decrease in the intensity of the signal is observed, since the beads are taken apart by the displacer. This allows the measurement of an IC_{50} value, defined as the concentration at which half of the displacement has occurred.

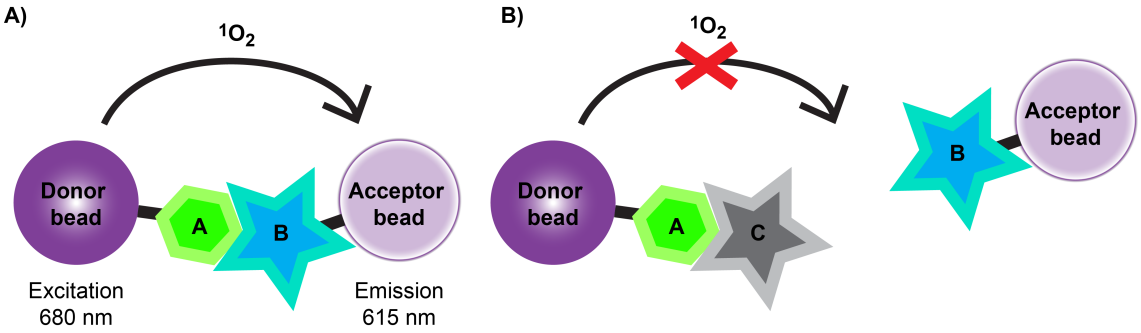


Figure 1.9 – AlphaLISA principles. A) Direct experiment. Upon interaction, **A** and **B** bring the beads together. The donor bead is excited with light at 680 nm and transfers singlet oxygen to the acceptor bead that emits luminescence at 615 nm. **B) Competition experiment.** In the presence of a competitor (**C**), **B** no longer binds to **A**, the beads are not close in proximity and there is reduced emission of light by the acceptor bead, resulting in a loss of signal.

1.5 Prior art

Prior work in the Ciulli laboratory has led to the identification of two novel binding sites in the trimeric subunit of the $CRL2^{VHL}$ complex constituted by pVHL, EloB and EloC (VBC) that are distinct from the HIF-1α binding site (Figure 1.10). These binding sites were identified in a fragment screening campaign that resulted in the crystal structures of three fragments in complex with VBC being solved (work by Dr. Inge van Molle, University of Cambridge). One of the binding sites was located on pVHL and the other one was located on EloC in an interface that, at the time, was thought to be involved in the interaction with Cul2 (Figure 1.9), based on the expected structure analogy with other CRL complexes for which structures were available then. This hypothesis was confirmed upon the release of the quaternary structure of $Cul2_{1–163}$–VBC (PDB: 4WQO)[82] in 2015, during the course of this project. Work is undergoing on the development and optimisation of these fragments by other members of the laboratory.

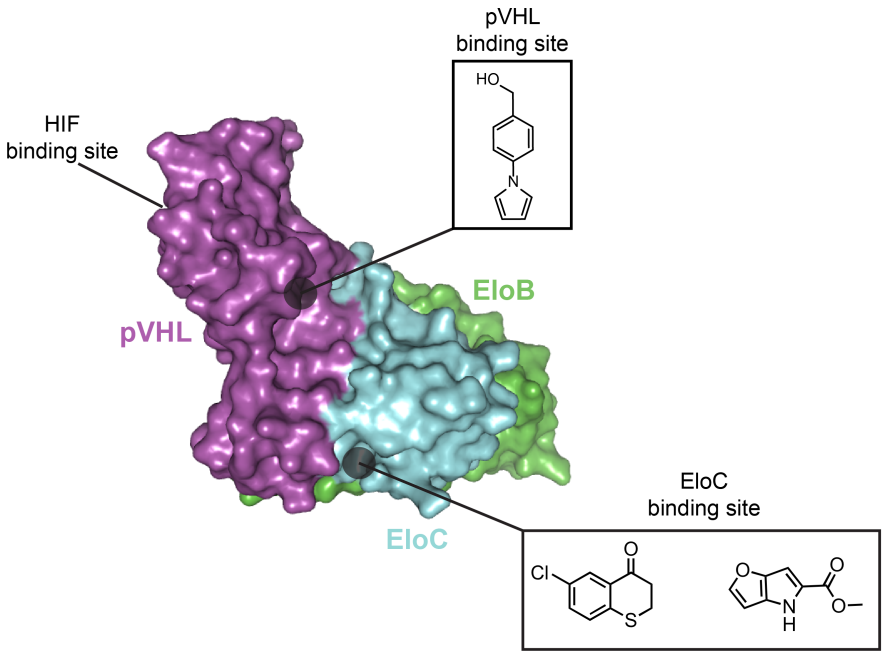


Figure 1.10 – VBC binding sites. Two new binding sites identified in the VBC subunit of the CRL2^{VHL} complex in a fragment screening campaign.

This encouraging discovery elucidated two potential ligandable sites, other than the HIF binding site, in VBC and motivated efforts of further understanding how targeting these sites, or other new binding sites, would influence the CRL2^{VHL} complex. The combination of the advantages of using peptides to target proteins with the previous success in the laboratory in developing potent modulators of the pVHL-HIF-1α interface designed from a peptide-bound structure reasoned the choice of peptides to probe these surfaces. Thus, a multi-disciplinary approach was devised to understand and chemically probe the CRL2^{VHL} complex, as model system of the multi-subunit Cullin RING E3 ligases family.

1.6 Aims

The aims of this project can be summarised in three main points:

1) To gain structural and biophysical insights into the $CRL2^{VHL}$ ligase complex and contribute to a better understanding of its assembly, interaction, dynamics and functioning as a protein complex.

2) To exploit the gained structural and biophysical knowledge as basis to design peptide-like chemical probes to explore the ligandability of the Cul2-VBC interface.

3) To probe protein surfaces within structural components of the VHL ligase with the ultimate goal of modulating protein-protein molecular recognition within the protein complex.

CHAPTER 2

STRUCTURAL AND BIOPHYSICAL INSIGHTS INTO THE $CRL2^{VHL}$ MULTI-SUBUNIT PROTEIN COMPLEX

2.1 Introduction

Cullin RING E3 ubiquitin ligases (CRLs) are the major subfamily of E3 ligases, enzymatic machineries that catalyse the transfer of ubiquitin from an E2 conjugating enzyme to a target substrate. Cullin is the core subunit of the CRL complexes and acts as a scaffold that brings to proximity the substrate and the E2 conjugating enzyme. The substrate is recruited by the receptor and the E2 enzyme loaded with ubiquitin is recruited by the RING finger protein.[83] All cullins present an elongated N-terminal domain (NTD) linked to a globular C-terminal domain (CTD).[18] The information available on this important family is limited by the existence of crystal structures of full length protein only of Cul1,[84] Cul4A[85] and Cul4B.[85] In addition, there are also structures for partially assembled components and there are several structures of portions of other Cullin proteins *i.e.* NTD or CTD alone or in complex with other subunits. However, the only structures available of whole CRL complexes are those of $CRL4A^{DDB2}$ (ref. 85), $CRL4B^{DDB2}$ (ref. 85), and $CRL1^{Skp2}$ (ref. 84). This lack of structural information restricts the full understanding of the mechanisms of activity and functioning of CRL multi-subunit molecular machines. Structural insights are critical also for drug discovery, as they can provide the basis for exploring PPIs with small molecules and modulate their function.[86] For instance, the structural knowledge of PPIs within certain CRLs has facilitated the ability to hijack E3 ligase complexes into recruiting non-natural selective substrates for ubiquitination and subsequent proteasomal degradation.[45,87,88]

The aim of the work presented in this chapter was to gain structural insights on the mechanisms of functioning of the $CRL2^{VHL}$ and to understand the driving forces of particular PPIs within the complex. It is described the use of X-ray crystallography to unveil the first structure of the $CRL2^{VHL}$ complex to

3.9 Å resolution and the characterisation of the Cul2–VBC interaction, using biophysical techniques such as AlphaLISA and ITC. Residues that are critical for the strength of the Cul2–VBC PPI were identified and the basis for the selectivity of $CRL2^{VHL}$ for Cul2 over Cul5 was also explored.

2.2 Results

2.2.1 Protein expression and purification

The $CRL2^{VHL}$ complex was assembled by mixing a trimeric subunit composed of pVHL, EloB and EloC (VBC) with a dimeric subunit composed of Cul2 and Rbx1. The choice to express the pentameric complex as two separate subunits was supported by the availability of protocols for the expression and purification of these components.

VBC was expressed in *E. coli* according to a previously published protocol.[52] Briefly, BL21(DE3) *E. coli* bacteria were co-transformed with a plasmid for the expression of pVHL with a hexahistidine tag at its N-terminus and a second plasmid for the expression of both EloB and EloC. Co-expression of these three proteins resulted in VBC complex formation that was then purified by affinity chromatography, followed by ion-exchange chromatography and, finally, by size-exclusion chromatography. During the purification process the hexahistidine tag was removed by cleavage with TEV protease. Following this protocol, VBC was purified routinely with an estimated purity above 95% and with a yield of about 15 milligrams per litre of culture.

The Rbx1–Cul2 heterodimer subunit was co-expressed in a baculovirus system in *Sf21* insect cells. The pFastBacDual plasmid containing Rbx1 and Cul2 with a N-terminal Dac tag (affinity tag based on penicillin-binding protein 5 that binds ampicillin)[89] was a gift from Dr. Mark Peggie (DSTT). This plasmid was transformed into DH10Bac *E. coli* cells in order to generate the bacmid and the colonies containing the bacmid were selected by blue/white screening. The isolated bacmid was then transfected into *Sf21* cells in adherent cell culture and the P0 virus was harvested after seven days of culture. The cell pellet was used to confirm the success of the transfection through assessing the expression of

Cul2 by immunoblot. The virus stock was amplified by infecting *Sf21* cells in suspension cell culture with the P0 virus. The cells were incubated for five days at 26.5 °C, with moderate shaking (130 rpm) in the dark and after this time the supernatant was collected (P1 virus) and stored at 4 °C in the dark. Protein expression consisted in infecting *Sf21* cells in suspension cell culture at a density of 1.5 million cells/ml with 1:100 P1 virus, followed by incubation of three days in the dark, at 26.5 °C and with moderate shaking (130 rpm). The amount of virus used for infection and the length of incubation were optimised in order to obtain the highest amount of protein. The Rbx1–Cul2 purification protocol was based on a protocol developed by Dr. Axel Knebel and Clare Johnson (MRC-PPU/DSTT), and it has been slightly optimised to improve the yield and quality of the protein by modifying lysis buffer and lysis method and by introducing a gel filtration step. After harvesting the cells by centrifugation and lysis by French press method, the Dac tag-containing Cul2 in complex with Rbx1 was isolated from the cell lysate using ampicillin-modified sepharose resin. The Dac tag was then cleaved with TEV protease while the complex was bound to the resin. Rbx1–Cul2 was recovered through filtration and the filtrate was purified through gel filtration chromatography, in order to isolate monomeric Rbx1–Cul2 (Figure 2.1). The yield varied from 0.4-2.5 milligrams of protein per litre of culture.

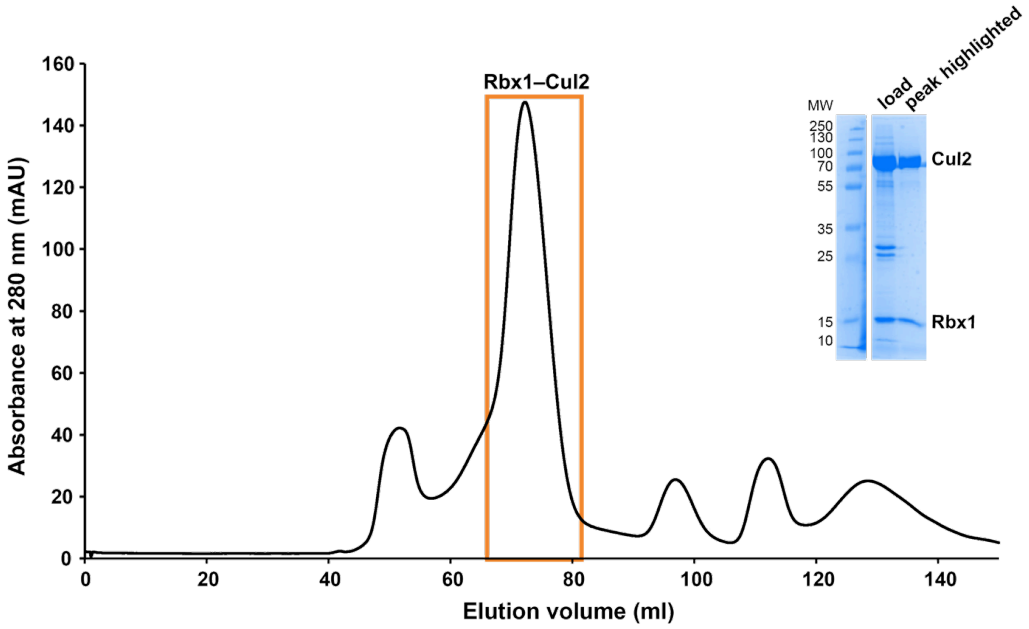


Figure 2.1 – Purification of Rbx1–Cul2 by size-exclusion chromatography. Chromatogram resulting from Rbx1–Cul2 purification on a HiLoad 16/600 Superdex 200 gel filtration column. The first peak is eluted in the void volume and corresponds to aggregated Rbx1–Cul2, the second peak corresponds to monomeric Rbx1–Cul2 and the following peaks correspond to lower molecular weight contaminants. The Rbx1-Cul2 sample is obtained with a high degree of purity, as confirmed by the SDS-PAGE.

2.2.2 Crystal structure of the CRL2^{VHL}

The CRL2^{VHL} complex composed by Cul2, EloBC, pVHL and Rbx1 was re-constituted by mixing VBC and Rbx1–Cul2, and purifying the pentameric complex by size exclusion chromatography (Figure 2.2). The appearance of a new peak eluting earlier than Rbx1–Cul2 suggested the formation of the pentameric Rbx1–Cul2–VBC complex. This was confirmed by SDS-PAGE that showed the presence of the five subunits under the same chromatographic peak. Separately, the same complex was prepared in the presence of a HIF-1α 19-mer peptide (residues 559-577) bound to pVHL, mimicking one of the natural substrates of the CRL2^{VHL} ubiquitin ligase.

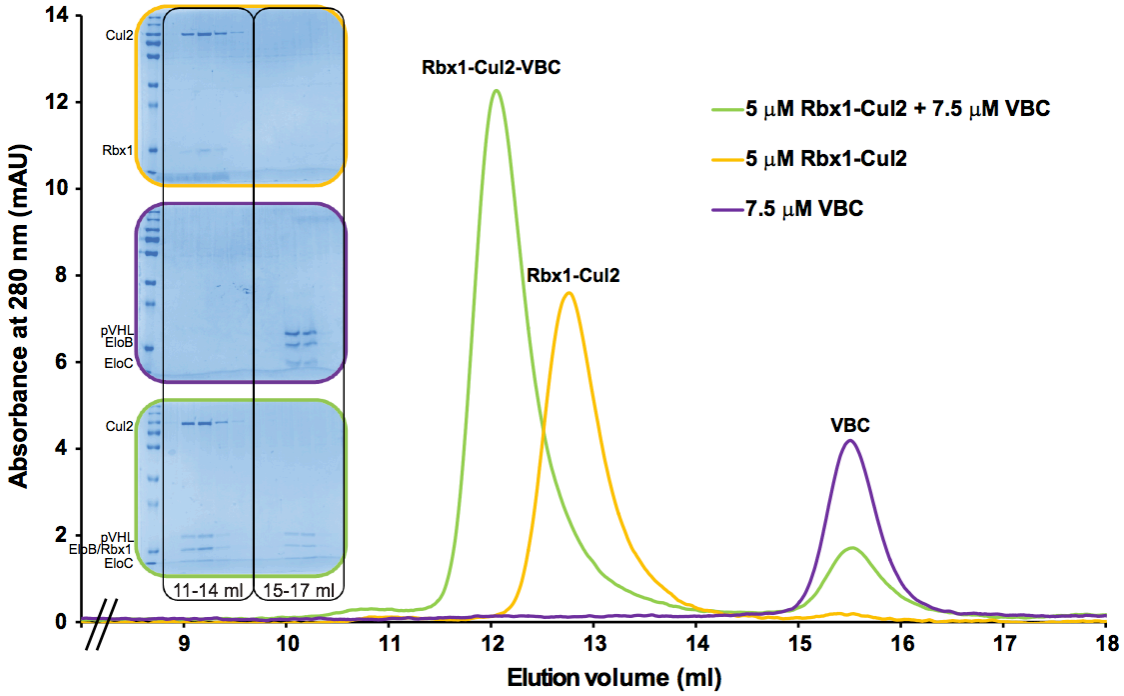


Figure 2.2 – Co-elution gel filtration of the CRL2^{VHL} complex. Elution profile from an analytical Superdex 200 Increase gel filtration column. 7.5 μM VBC (purple), 5 μM Rbx1–Cul2 (yellow) and a mixture of 7.5 μM VBC and 5 μM Rbx1–Cul2 (green) and corresponding SDS-PAGE. The elution profile of the mixture of VBC and Rbx1–Cul2 supported by the SDS-PAGE confirms the formation of a pentameric complex.

The sample was concentrated and used to set up crystallisation screenings (JSCG+ from Qiagen and ProPlex from Molecular Dimensions) of both complex forms (with and without the HIF-1α 19-mer peptide). Hit conditions were identified for the HIF-1α containing complex in the two screenings. The hits were reproduced and one of the initial conditions was followed up and optimised by grid screening, additive screening and introduction of streak seeding. In the final condition, promising diffracting crystals were obtained at 20 °C within two days, in a crystallisation solution containing Tris pH 7.6, ammonium sulphate, PEG4000 and 1,4-dioxane or acetonitrile (Figure 2.3). Despite diffracting X-rays, the crystals obtained were, to some extent, elastic and fragile. A number of cryoprotectant solutions were tested in an attempt to improve the quality of the data collected and 20% (v/v)

2-methyl-2,4-pentanediol (MPD) in the mother liquor was identified as the most adequate cryoprotectant based on the quality of the diffraction. The dataset for crystals of the $CRL2^{VHL}$ complex that led to solving the structure of the complex was collected at the Diamond Light Source synchrotron and the crystal diffracted X-rays up to 3.9 Å. Despite numerous efforts to improve the quality of the crystals and of the datasets collected, this was the best result obtained. The Matthews coefficient[90] calculation retrieved a solvent percentage of 63.9%, consistent with the presence of one protomer per asymmetric unit arranged in a $C222_1$ space group (Table 2.1).

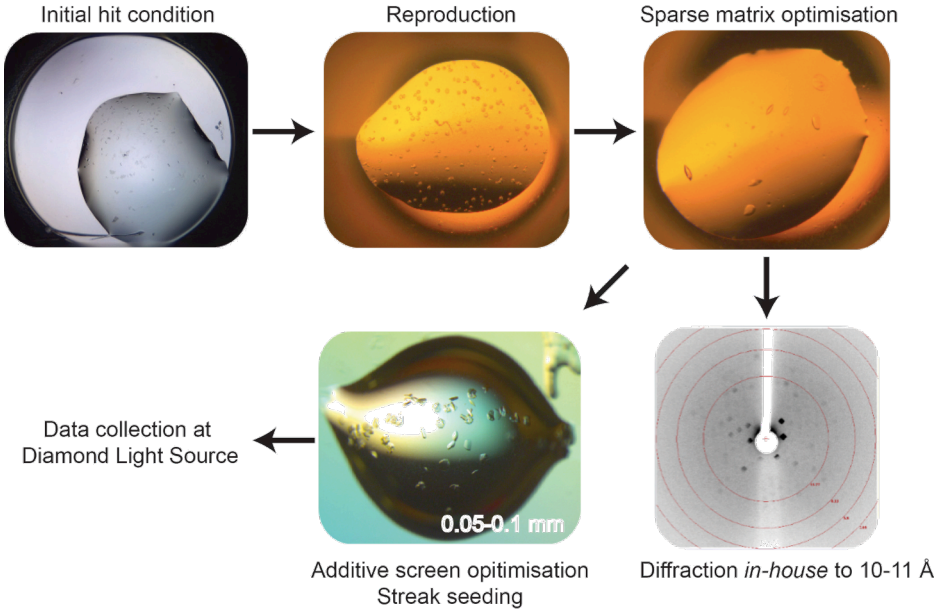


Figure 2.3 – Optimisation of crystallisation conditions. Process of optimisation of the $CRL2^{VHL}$ crystals from the initial hit condition to the crystals used for data collection.

The structure was solved at 3.9 Å using a combination of molecular replacement (MR) and iterative rounds of model building and refinement. An homology model of Cul2 (obtained through Chimera)[91] based on the structures of Cul1, Cul4, $Cul5_{NTD}$ and $Cul5_{CTD}$ (PDB's 1U6G,[92] 4A0K,[85] 4JGH,[93] 3DPL,[19] respectively) was used as template model to find a molecular replacement

solution for the structure of Cul2 in MrBUMP.[94] Next, EloBC dimer was placed by MR in Phaser,[95] using as a template the corresponding subunits from a VBC–HIF-1α structure (PDB 4AJY)[52]. By then, in the region corresponding to the Rbx1 N-terminal tail (residues 17-35) clear unmodelled electron density could be observed. This region is conserved in all Cullin-bound Rbx1 structures available in the PDB, allowing its correct positioning into the structure. This was followed by the identification of the RING domain of Rbx1 by MR using as template an existing Rbx1 structure (PDB 2LGV)[96] in Phaser. Other models of the Rbx1 RING domain based on structures in complex with other Cullins were also tried as templates for MR, but yielded no or unsatisfactory results, in which the solutions were not appropriately fitting the electron density. The area where pVHL was expected to be found, according to the previous structure of VBC in complex with $Cul2_{1\text{-}163}$,[82] was, at this point, filled with patches of positive electron density. However, fitting in a complete model of pVHL proved quite challenging and despite various attempts at molecular replacement using a diversity of template models, no reasonable solution was found. Nonetheless, the electron density corresponding to the three helix cluster structure of the VHL-box[97] was unambiguous and allowed the manual fitting of this region that is conserved in all the available crystal structures. Unfortunately, for the remaining residues (54-156) constituting the substrate-binding domain of pVHL, and for the 19-mer HIF-1α peptide no electron density was observed.

Table 2.1 – Data collection and refinement statistics. Statistics in parentheses indicate those for the highest resolution shell.

Data collection	
Wavelength (Å)	0.9282
Space group	$C222_1$
Cell dimensions	
a, b, c (Å)	86.0, 191.0, 238.9
α, β, γ (°)	90, 90, 90
Molecules/ASU	1
Resolution (Å)	95.48 - 3.90 (4.27 – 3.90)
R_{merge} (%)	11.4 (73.2)
$I/\sigma(I)$	10.2 (2.2)
Completeness (%)	100 (100)
Redundancy	5.7 (5.3)
CC (1/2)	0.857 (0.798)

Refinement	
Resolution (Å)	95.48 - 3.9
Unique reflections	18326
R_{work}/R_{free} (%)	30.15/34.61
Average B factor (Å^2)	191
No. non-hydrogen atoms	7734
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.692
Ramachandran analysis	
Preferred regions (%)	90.65
Allowed regions (%)	9.13
Outliers (%)	0.22

The above described steps were intercalated with rounds of refinement in Refmac5[98] and Phenix[99] and map sharpening by negative B-factor correction was performed to improve the quality of the electron density maps and facilitate successful model building (Figure 2.4B).[100] Correctness and quality of the MR solutions were assessed by the MR score and by visual inspection of the solution, considering the adequate fitting to the electron density and packing of

the structure. Throughout the molecular replacement, docking and refinement rounds, the R_{work} and R_{free} values were considered to evaluate the appropriateness of the resulting models.

The $CRL2^{VHL}$ structure forms extensive crystal contacts with eight symmetry related molecules. The crystal contacts' surface extends through an area of ~ 3300 $Å^2$, which corresponds to about 60% of the surface area of the whole complex (~ 5200 $Å^2$). Cul2 is contacting four Cul2 symmetry molecules and two Rbx1 subunits from two of these symmetric molecules, establishing 21 hydrogen bonds and 5 salt bridges and accounting for the most extensive area of crystal contacts. The VHL box packs against another VHL box from a symmetric molecule, while the EloBC subunit is not extensively involved in any crystal contacts. Lastly, the Rbx1 RING domain establishes mainly hydrophobic crystal contacts with Rbx1–Cul2 from another asymmetric unit (ASU). These contacts extend through an area of ~ 360 $Å^2$, which corresponds to about 17% of the total area of interface of Rbx1 with Cul2 (~ 2100 $Å^2$).

The final model (PDB: 5N4W) is constituted by the full-length Cul2 binding Rbx1 at the C-terminal domain and the VHL-box-EloBC at the N-terminal domain, revealing for the first time the whole Cul2 structure and the interface between Rbx1 and Cul2.

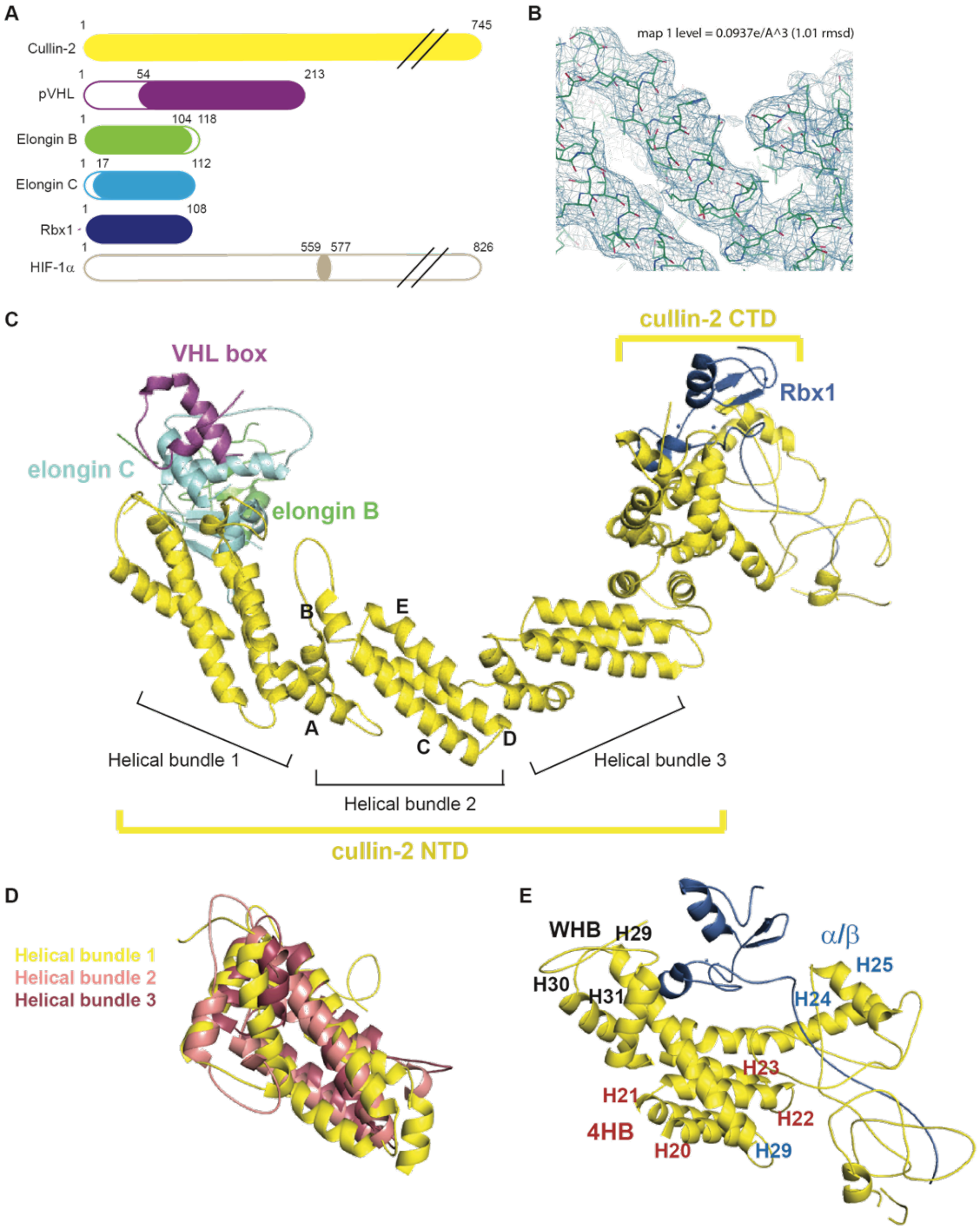


Figure 2.4 – Crystal structure of the CRL2^{VHL} complex. A) Diagrams of the primary structure of each of the protein subunits, illustrating the constructs (coloured bars) used to obtain the crystal structure. **B)** Representation of the $2F_o–F_c$ electron density contoured at 1σ over a portion of the model with the backbone and side chains shown in green. **C)** Crystal structure of the CRL2^{VHL} where each of the protein subunits are identified; Cul2 is divided in N-terminal (NTD) and C-terminal (CTD) domain and the NTD is composed of three helical bundles, constituted by five α-helices (A-E). **D)** The three helical bundles of the NTD are superimposable with a maximum RMSD of 4.5 Å over the Cα. **E)** The CTD of Cul2 is organised in a four-helical bundle (4HB) motif (helices H20-H23), an α/β domain (helices H24, H25 and H29) and a winged-helix motif, WH-B (helices H29-H31). Rbx1 binds to the CTD and functions as recruiter of the E2 conjugating enzyme to the complex.

2.2.2.1 Cullin-2 structure highlights its inherent flexibility

The structure of Cul2 in the $CRL2^{VHL}$ reveals a similar architecture to structures of other Cullins. It is characterised by the classical elongated shape divided in two main domains: the Cullin NTD that recruits the substrate-binding domain through an adaptor subunit and the CTD, a globular domain that binds Rbx1, which in turn recruits the E2 conjugating enzyme to the complex. Residues 1-384 are comprised in the NTD and arranged in three helical bundles (also called Cullin repeats) each composed of five α-helices. These three Cullin repeats superpose with a root mean square deviation (RMSD) for the Cα atoms of up to 4.5 Å (Figure 2.4D). The CTD (residues 385-745), at the other end, is organised in a four-helical bundle (4HB) linked to the NTD, an α/β domain and a winged-helix motif (WH-B) – Figure 2.4E. Some residues at the CTD (625-634 and 645-660) were disordered and could not be modelled in the structure.

Cul2 was one of the few members of the Cullin-family lacking structural information on the full NTD. By superposing all the existing structures of complete Cullin NTDs the flexibility of this domain becomes evident (Figure 2.5). It was observed that all the NTDs align most closely through the second Cullin repeat, showing considerable differences in their first and third helical bundles' relative position, which is consistent with intra-domain bending motion (Figure 2.5A). Additionally, the structural alignment of full-length Cul1, Cul2 and Cul4A highlights an interesting difference on how the CTD is packed against the NTD in each of the structures (Figure 2.5B-C). The globular domain in Cul2 is substantially moved towards the NTD in regards to other Cullins. The distances between the very top residue at the CTD (the Lys that gets modified with NEDD8 *i.e.* Lys689 in Cul2, Lys720 in Cul1 and Lys705 in Cul4) and the

top residue in helix α1 of the NTD (Phe10 in Cul2, Leu17 in Cul1 and Thr62 in Cul4) are at least 10 Å shorter in Cul2 (100 Å for Cul2, 110 Å for Cul1 and 113 Å for Cul4). $Cul1_{CTD}$ is rotated by 29° compared to $Cul2_{CTD}$, whereas $Cul4_{CTD}$ presents an angle of rotation of about 33° compared to $Cul2_{CTD}$.

Flexibility of Cullins is deemed essential for the mechanism of polyubiquitination of substrates.[101] The structural analysis of Cul2 underpins the previously reported flexibility of the scaffold and suggests that non-neddylated Cul2, alike the other non-neddylated Cullins, presents a closed state conformation.

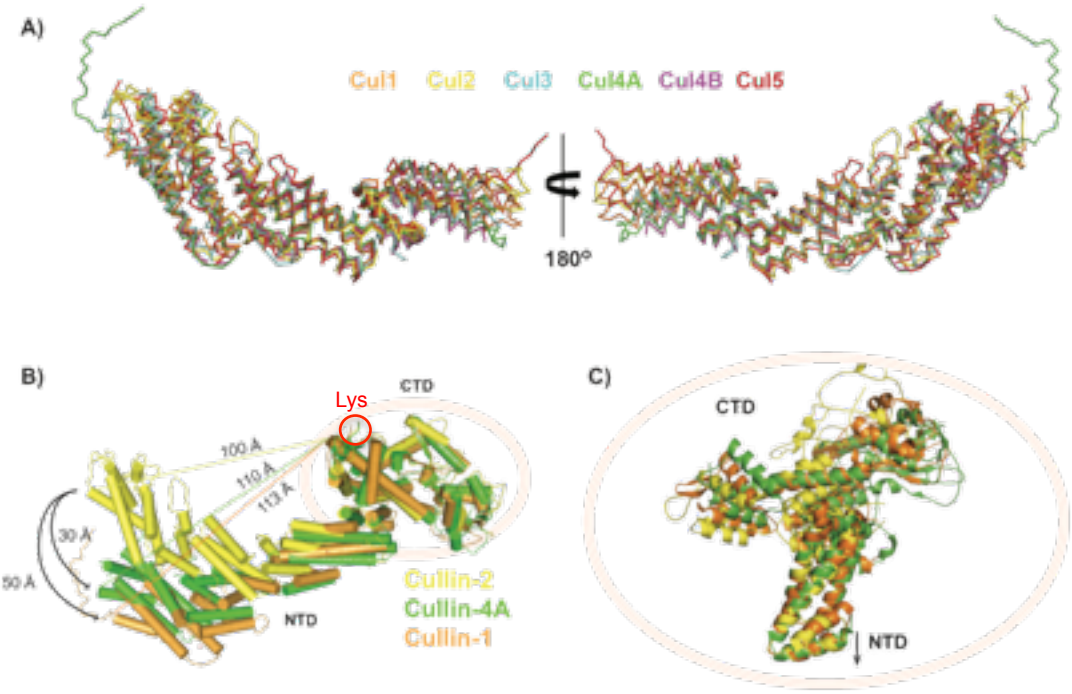


Figure 2.5 – The $CRL2^{VHL}$ complex is dynamic and flexible. A) Superposition of the crystal structures of the full NTD of Cul1 (PDB 1U6G),[92] Cul2 (PDB 5N4W), Cul3 (PDB 4HXI),[102] Cul4A (PDB 4A0K),[85] Cul4B (PDB 4A64) and Cul5 (PDB 4JGH)[93] illustrating the existence of hinge points in the linkers between cullin repeats. **B)** Full length structures of Cul2, Cul4A and Cul1 superimposed by the CTDs reveal considerable inter-domain flexibility through a hinge point between NTD and CTD. In the Cul2 structure the CTD is considerably closer to the NTD than in other cases. **C)** Close up view of the CTDs of Cul1, Cul2 and Cul4A when the proteins are aligned by the third cullin repeat of the NTD illustrating the different relative orientations of the CTDs in the full-length structures.

2.2.2.2 Rbx1 presents a new orientation

The Rbx1 structure is composed of an N-terminal tail arranged in a long β-strand engaging the Cullin CTD and a variant RING finger domain. The variant RING presents an extended region containing a third zinc ion coordinated by three cysteines and one histidine, in addition to the canonical region containing two zinc ions.[84] The structure and position of the N-terminal tail of Rbx1 in the complex are highly conserved amongst different CRL structures. Van der Waals contacts mainly established by Phe22 and Trp27 anchor the β-strand of Rbx1 to Cullin CTD. On the contrary, the RING domain is flexible and it has been found in a number of different orientations, even in different molecules within the same asymmetric unit of a previously solved crystal structure.[19] Interestingly, the conformation of Rbx1 in the structure presented here was unprecedented. The observation of minor crystal contacts formed by Rbx1, as described in section 2.2.2, supports that the observed conformation of Rbx1 captures a significantly populated conformation in solution. It sits between Rbx1-Cul5~NEDD8 and Rbx1-Cul5 conformations resembling a transient state *en route* amid the two (Figure 2.6). The flexibility associated with the RING domain has been studied before with the intent of elucidating the structural and mechanistic details of the catalytic activity of RING ubiquitin ligases.[19,84,92,103-105] It has been demonstrated that the RING domain dramatically re-orients its position relative to Cullin upon neddylation of Cul5.[106] This conformational change could be considered as an adaptation of the enzyme prior to ubiquitination. The area of interface between Rbx1 and Cul5 (1707 $Å^2$), in this so-called open conformation, is contributed solely by the N-terminal tail (Table 2.2). In fact, the total areas of interface between Rbx1 and the different Cullin subunits differ in regards to the orientation of the Rbx1 RING

domain (Table 2.2). In this new crystal structure, the linker between the N-terminal tail and the RING domain, albeit not completely extended, is not fully retracted either and the Rbx1–Cul2 PPI covers an area of about 2100 Å^2. Such interface area is greater than in the Rbx1–Cul5~N8 and in the Glomulin–Rbx1–Cul1 complexes but smaller in comparison to any of the other structures (Table 2.2).

These findings suggest a probable trajectory of the Rbx1 RING domain from the non-neddylated to neddylated forms of the complex, *en route* to the fully active E3 ligase.

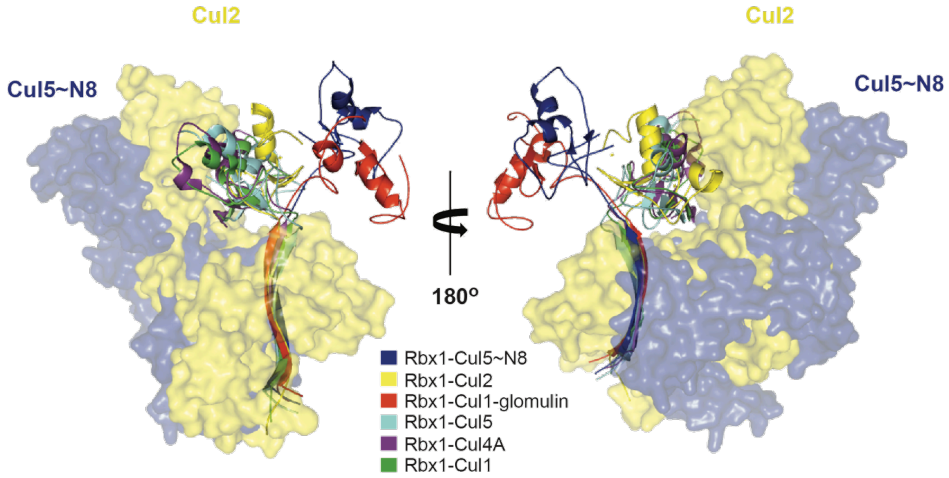


Figure 2.6 – Rbx1 presents a new orientation. Superposition of the CTDs of six cullins - Cul1 (PDB: 1U6G),[92] Cul2 (PDB: 5N4W), Cul4 (PDB: 4A0C),[85] Cul5 (PDB: 3DPL),[19] Cul5~NEDD8 (PDB: 3DQV),[19] Glomulin-Rbx1-Cul1 (PDB: 4F52)[107] - complexed with Rbx1. The new structure of Rbx1 in complex with Cul2 unveils a novel orientation of its RING domain, resembling an *en route* conformation between Rbx1-Cul5 and Rbx1–Cul5~NEDD8.

Table 2.2 – Different areas of interface in different Rbx1–Cullin complexes. The areas of interface were calculated (PISA) as well as the rotation angle and shift along the rotation axis (Chimera), and compared to previously published structures.

Interaction pair	Area of interface (Å^2)	Rotation angle (º)	Shift along the axis (Å)	PDB
Rbx1-Cul5~NEDD8	1707	124	10.2	3DQV
Rbx1-Cul2	2122	-	-	5N4W
Rbx1-Cul5	2307	118	-1.1	3DPL
Rbx1-Cul4	2708	118	-1.9	4A0C
Rbx1-Cul1	3396	114	-3.4	1U6G
Glomulin-Rbx1-Cul1	1718	138	17.1	4F52

2.2.2.3 The interface between Cul2 and VBC

The Cul2–VBC interface is spread across three contact regions (Figure 2.7A). The first comprises the N-terminal loop of Cul2, which assumes a role in stabilising the interaction by establishing critical hydrophobic contacts (Figure 2.7B), with Leu3 enclosed in a hydrophobic pocket on EloC (delineated by Met105 and Phe109) and Pro5 involved in contacts with Val181 in pVHL and Met105 in EloC (Figure 2.7B). The importance of residues belonging to the N-terminal loop in this PPI is observed in Cul2 and Cul3.[102] The second region of interaction involves charge complementarity at the interface between helix α5 of Cul2 and the substrate receptor, with Lys159 (pVHL) forming an hydrogen bond with Gln111 (Cul2) and Asp187 (pVHL) establishing a salt bridge with Lys114 (Cul2, Figure 2.7C). These electrostatic contacts were thought to be responsible for the selective recruitment of Cul2 over Cul5 to this complex.[82] In the last area of interaction, helix α2 of Cul2 is packed against pVHL and EloC surfaces, with residues Asn36, Phe39, Tyr43 and Val47 contributing to form a mainly hydrophobic interface (Figure 2.7D). There is no direct interaction between EloB and Cul2. The total buried surface areas of the Cul2–VBC PPI and the orientation of the VHL-box-EloBC relative to Cul2 are in agreement with those observed in the quaternary structure of $Cul2_{NTD}$ (Figure 2.8 and Table 2.3). Moreover, superposition of the two structures results in an RMSD of 0.866 Å for the Cα atoms, suggesting a structurally rigid and conserved interface.

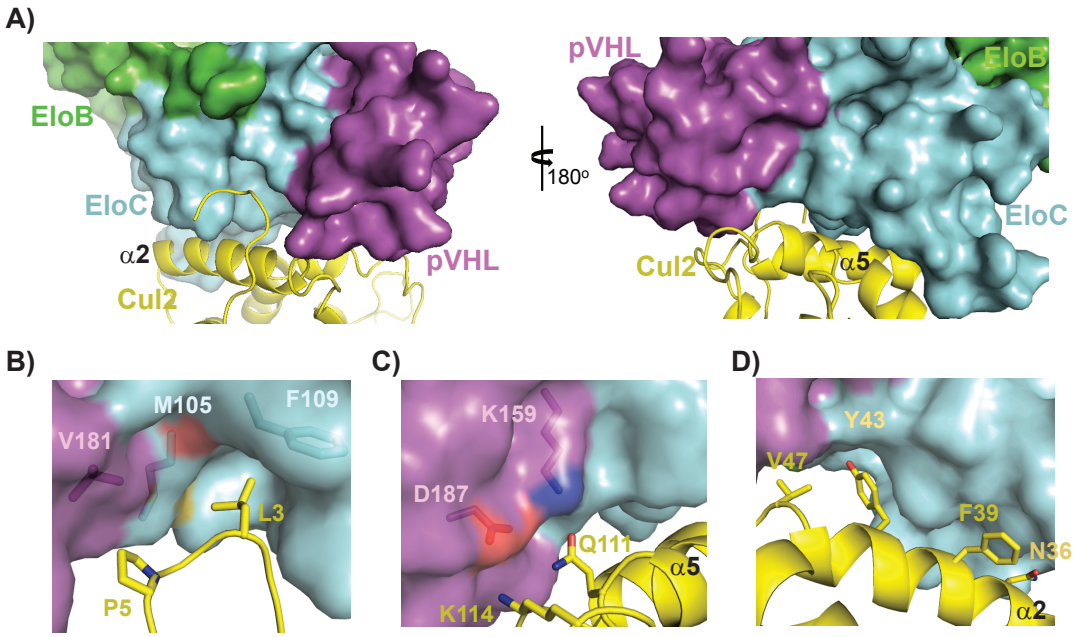


Figure 2.7 – The Cul2–VBC interface. A) Contacts between $Cul2_{NTD}$ and VBC. As in similar CRL complexes, helix α2 of Cullin is the closest to the adaptor and substrate receptor subunits, establishing, predominantly, hydrophobic contacts. **B)** The N-terminal tail of Cul2 plays a substantial role in the PPI. Leu3 is accommodated in a hydrophobic pocket on the surface of EloC and Pro5 is involved in a three-way interaction between Val181 from pVHL and Met105 from EloC. **C)** Residues Asp187 and Lys159 in pVHL establish an electrostatic network with residues Lys114 and Gln111 in the helix α5 of Cul2. **D)** Residues Asn36, Phe39, Tyr43 and Val47 participate in hydrophobic contacts at the interface between helix α2 and EloC and pVHL.

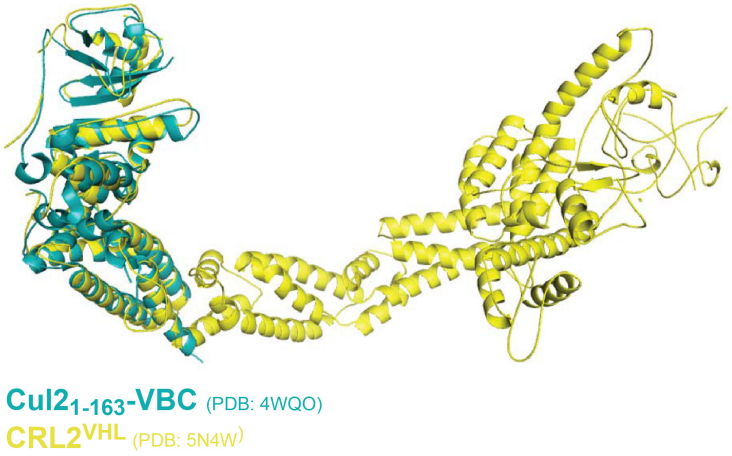


Figure 2.8 – Superposition of the structures of $Cul2_{1-163}$–VBC and $CRL2^{VHL}$. The structure of the $CRL2^{VHL}$ complex (yellow) is superposed with the quaternary structure of $Cul2_{1-163}$–VBC (cyan). The structures align with and RMSD of 0.866 Å over the Cα atoms.

Table 2.3 – Comparison of the areas of interface between Cul2 and VBC in the two structures available.

	Buried surface area involving Cul2 interfaces (Å^2)		
PDB	pVHL-Cul2	EloC-Cul2	Total
4WQO	322	944	1270
5N4W	328	1022	1350

2.2.3 Neddylation of the CRL2^{VHL} complex

It is well established that the attachment of a ubiquitin-like molecule, NEDD8, to the CTD of Cullin is required for activation of the ligase activity.[23,108-111] It is not clear, however, whether the neddylation happens before of after substrate recruitment. To assess this, an *in vitro* neddylation assay[112] was performed in the presence and in the absence of the HIF-1α 19-mer peptide. The E1 and E2 reagents and the protocol were provided by Dr. Axel Knebel (MRC-PPU/DSTT). In this experiment, two E2 conjugating enzymes were tested, UBE2M and UBE2F. The neddylation reaction was incubated for two hours at 37 °C and the samples were then analysed by SDS-PAGE (Figure 2.9). The bands corresponding to Cul2 were confirmed by immunoblot.

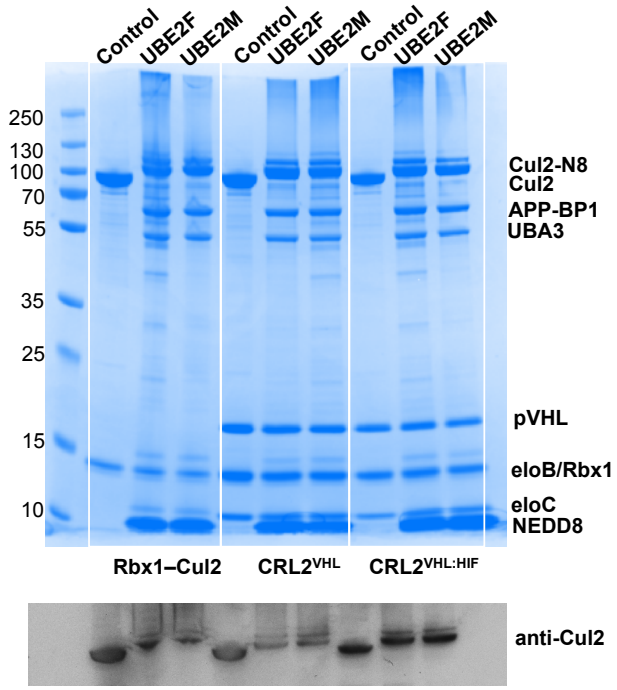


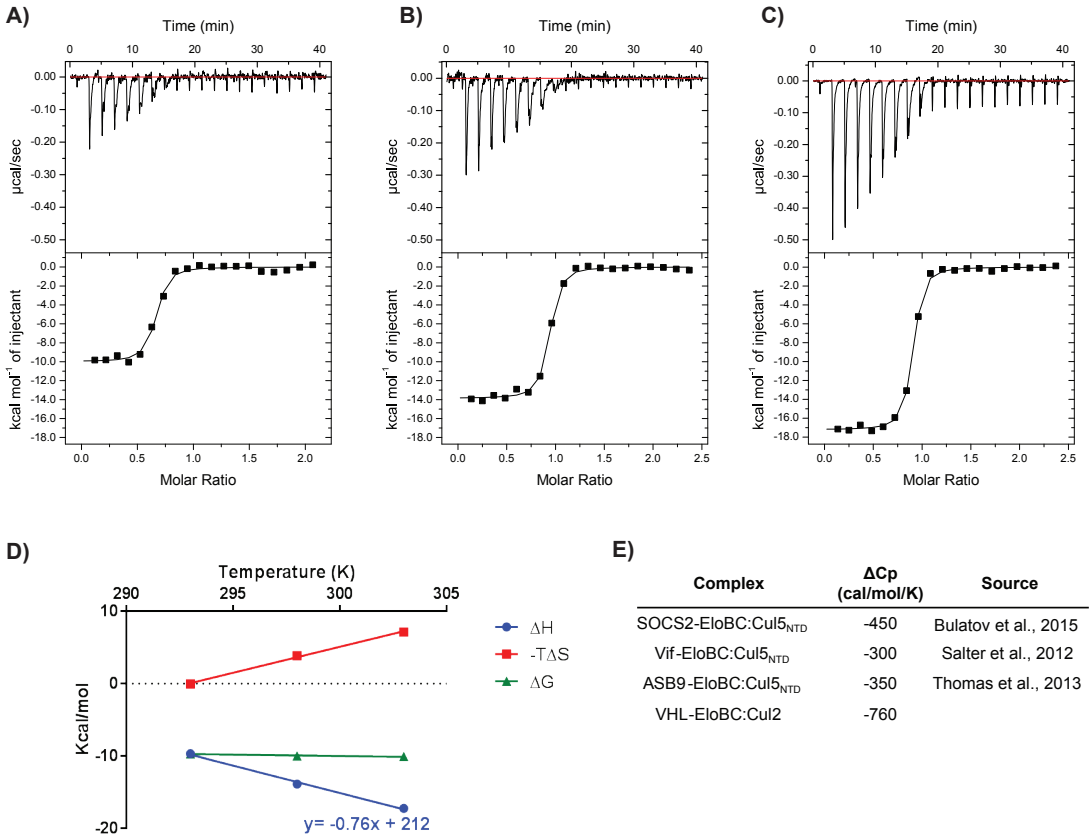
Figure 2.9 – Neddylation assay. A neddylation assay was performed with Rbx1–Cul2, $CRL2^{VHL}$ and $CRL2^{VHL:HIF}$ as substrates to evaluate the importance of the presence of the CRL2 substrate for the reaction to occur. Two E2 conjugating enzymes were tested, UBE2M and UBE2F. UBA3-APP-BP1 corresponds to the E1 enzyme.

The results showed that all the substrates for neddylation *i.e.* Rbx1–Cul2, $CRL2^{VHL}$ and $CRL2^{VHL:HIF}$ were converted to the neddylated form in the presence of either of the two E2 enzymes tested. Although unlikely to happen *in vivo*, Cul2 is neddylated in the absence of VBC *in vitro*.[113] Having that in mind, in this *in vitro* assay, no difference was observed in the neddylation profile of the $CRL2^{VHL}$ complex in the presence or in the absence of the substrate-like peptide.

2.2.4 Biophysical characterisation of the Cul2–VBC interaction

The lack of biophysical data regarding PPIs within the $CRL2^{VHL}$ complex motivated an in-depth characterisation of specific interactions. Despite persistent efforts to obtain Cul2 alone as functional protein it was not possible to achieve soluble, monomeric protein that could form a complex with VBC

(discussed in Chapter 3). As Cul2 and Rbx1 had to be co-expressed to obtain soluble and functional protein, the investigation of the PPI between these two subunits was prevented, and, therefore, it was decided to focus on the PPI between Cul2 and VBC. ITC experiments were performed, titrating VBC into Rbx1–Cul2. Expectedly strong binding affinity (K_d = 42 nM) and large exothermic binding enthalpy (ΔH = – 17.2 kcal x mol^{-1}) at 303 K were measured (Figure 2.10C). As a way of comparison, data already available on Cul5–EloBC complexes, namely Cul5$_{NTD}$–ASB9–EloBC and Cul5$_{NTD}$–SOCS2–EloBC, reported binding affinities of 140–220 nM[114,115] and 7–47 nM,[116-118] and binding enthalpies of ΔH = – 8.3 kcal x mol^{-1} and ΔH = – 4.8 kcal x mol^{-1}, respectively, for the corresponding interactions. The temperature-dependence of the binding parameters was also investigated and a change in heat capacity ΔC_p = – 760 cal x mol x K^{-1} was determined (Figure 2.10D). The change in heat capacity is a thermodynamic parameter defined as the change in energy (heat) with temperature ($\Delta C_p = \Delta H/\Delta T$).[119] In biological systems, negative ΔC_p values are often observed, meaning that an interaction becomes more exothermic with the increase of temperature. ΔC_p values determined for other similar complexes are significantly less negative than the one obtained for the Cul2–VBC system (Figure 2.10E), which suggests a greater buried surface area upon Cul2–VBC interaction, leading to a more stable complex. Consistent with this hypothesis, based on the crystal structures available, the area buried upon interaction in Cul2–VBC is 2621 Å^2, which is much greater than in Cul5–SBC (1945 Å^2) and in Cul5–VifCβFBC (1605 Å^2, Table 2.4).



E)

Complex	ΔCp (cal/mol/K)	Source
SOCS2-EloBC:Cul5$_{NTD}$	-450	Bulatov et al., 2015
Vif-EloBC:Cul5$_{NTD}$	-300	Salter et al., 2012
ASB9-EloBC:Cul5$_{NTD}$	-350	Thomas et al., 2013
VHL-EloBC:Cul2	-760	

Figure 2.10 - Biophysical characterisation of the interaction between VBC and Cul2. Isothermal titration calorimetry (ITC) data where VBC (200 μM) was titrated into Rbx1−Cul2 (20 μM) at **A)** 293 K, **B)** 298 K and **C)** 303 K. **D)** Temperature dependency of the thermodynamic parameters – ΔH, ΔG and -TΔS. The change in heating capacity, ΔC_p, was derived from the change in enthalpy with the temperature. **E)** Comparison of ΔC_p values for similar CRL systems.

The hydration of polar and apolar groups has been shown to promote changes in the heat capacity and this relation has been modelled by the equation: $\Delta C_p^{\text{hydration}} = C_a \, \Delta SASA_a + C_p \, \Delta SASA_p$, where SASA is the solvent-accessible surface area (apolar and polar) buried upon interaction and C_a and C_p are constants empirically determined representing the area coefficient per Å^2 contribution of residues in heat capacity change. Served by the newly solved crystal structure, the theoretical ΔC_p values were calculated using the different constants C_a and C_p (Table 2.5), which differ according to the datasets used to determine them, for example, the size and nature of the sample (proteins or cyclic dipeptides).[119-124] Interestingly, it was observed that

the experimental ΔC_p was larger than the theoretical values in most of the cases. This could indicate a conformational re-arrangement in solution upon binding that is not fully accounted for in the theoretical calculations.

Together, the biophysical and structural data suggest an extensive and tight PPI between the adaptor-receptor subunits and the scaffold in the $CRL2^{VHL}$ complex.

Table 2.4 - Buried surface area in the cullin interface with the receptor domains in three CRL complexes. PDBs 5N4W, 4JGH and 4N9F for $CRL2^{VHL}$, $CRL5^{SOCS2}$ and $CRL5^{Vif}$, respectively. The areas were calculated with GetArea.

	Area (Å^2)		
	Polar	Apolar	Total
VBC	5556	7915	13471
Cul2	17339	26443	43782
Cul2-VBC	21997	32635	54632
ΔSASA	-899	-1723	-2621
SBC	7270	10909	18179
$Cul5_{NTD}$	7710	11709	19418
$Cul5_{NTD}$-SBC	14361	21291	35652
ΔSASA	-619	-1326	-1945
VifCBFBC	10068	14102	24170
$Cul5_{NTD}$	6540	9737	16278
$Cul5_{NTD}$-VifCBFBC	16085	22757	38842
ΔSASA	-523	-1082	-1605

Table 2.5 - Comparison between experimental and theoretical ΔCp values[119-123,125] for the Cul2-VBC interaction.

	Source	ΔCp (cal/mol/K)
Experimental		-760
Theoretical	Spolar et al.	-425
	Murphy & Friere	-540
	Myers et al.	-400
	Makhatadze & Privalov	-692
	Robertson & Murphy	-384

2.2.5 Critical residues for the Cul2–VBC interaction

Identification of hotspots and elucidation of the main drivers of the tight PPI between Cul2 and VBC was the next step of investigation. Guided by structural analysis, the individual contributions of potential key residues were

quantified. This was accomplished by mutating contact residues involved in the PPI and assessing changes in binding affinities between mutants and wild-type proteins. The mutations were pointed at the three main regions of the Cul2–VBC interface, namely 1) the EloC pocket and adjacent area; 2) the helix α5-pVHL charge complementarity region; and 3) the hydrophobic contacts between helix α2 and EloC.

In the first region of the Cul2–VBC interface mutations were designed to evaluate the importance of the hydrophobicity and bulkiness of the side chain filling the EloC pocket. Leu3 in Cul2 was mutated to alanine and also to glycine to allow direct comparison with previous work.[82] Additionally, Pro5, a residue considered to potentially play a role in stabilising the conformation of the N-terminal loop of Cul2, which seemed important to direct the interaction, was mutated to alanine. Besides, Pro5 is also involved in hydrophobic contacts with Val181 (pVHL) and Met105 (EloC). In the other subunits, Val181 in pVHL was mutated to glycine and Met105 and Phe109 in EloC were mutated to alanine, addressing the consequences of disrupting the EloC hydrophobic pocket. In the second region, to understand the importance of electrostatic interactions in the complex formation, Gln111 and Lys114 in helix α5 of Cul2 were mutated to leucine and glutamate, respectively, and Lys159 and Asp187 in pVHL were mutated to glutamate and lysine, respectively. Lastly, in the third region of the interface residues Asn36, Phe39, Tyr43, Val47 on Cul2 were mutated to alanine, as those residues are involved in hydrophobic contacts with pVHL and EloC surfaces.

2.2.5.1 Site-directed mutagenesis, protein expression and purification

The mutations were performed by site-directed mutagenesis and the primers used are reported in Chapter 6. DNA sequencing of the plasmids confirmed the correct identity of each mutation. VBC mutant proteins were expressed and purified according to the protocol already described. For the Rbx1–Cul2 mutants, virus stocks for the mutant proteins were generated and then used for protein expression; the purification protocol used was the same as that which was previously described. All the mutant proteins were analysed by mass spectrometry to confirm the correct change in molecular weight for the corresponding mutation.

2.2.5.2 AlphaLISA experiments

All the mutant proteins generated were initially tested in an AphaLISA[126] assay to determine IC_{50} values from dose-response curves. A competition experiment was designed, to assess the ability of a competitor (protein with mutation) to disrupt the interaction between the two native protein subunits. VBC and Rbx1–Cul2 wild type were also titrated for reference. According to this experiment, all the mutations were disruptive to the Cul2–VBC interaction (Figure 2.11), although the extent of the disruption varied greatly (the raw data can be consulted in the Appendix I). It was observed that the most destructive mutations *i.e.* the ones leading to a larger increase in the IC_{50} values were mostly of residues located in the EloC pocket area. Particularly, the mutation of Met105 and Phe109 increased the IC_{50} by 75 and 55-fold, respectively. These results also showed that while Q111L and D187K mutations had a considerably disruptive effect, K114E and K159E mutations in the helix α5 interface had a

modest effect in the binding affinity. The mutations of residues in helix α2 interface led to an increase in the IC_{50} values by about 3-fold.

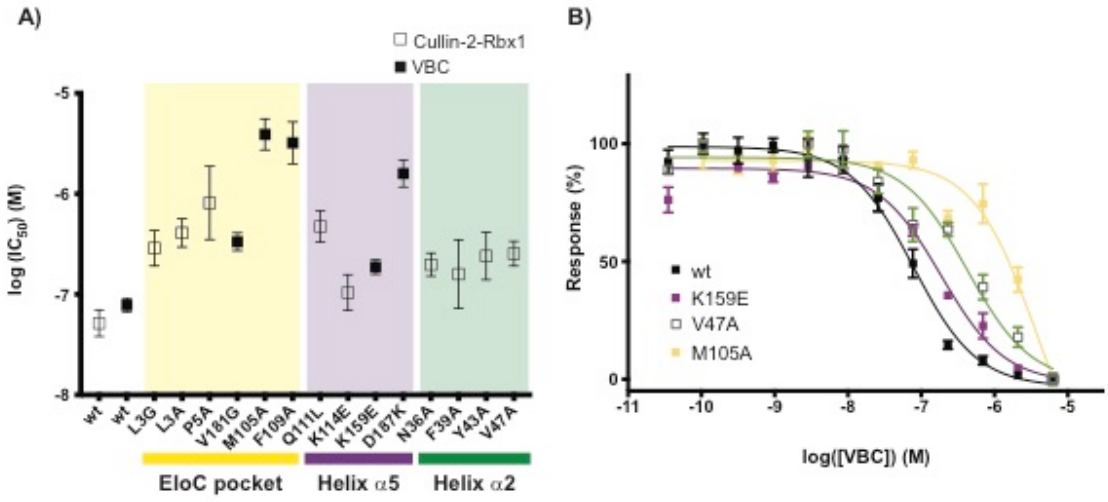


Figure 2.11 - Residues in the EloC hydrophobic pocket reveal critical for the strong binding affinity of Cul2–VBC. A) Plot of the log(IC_{50}) values resulting from the AlphaLISA competition experiment where mutant proteins of VBC or Cul2 were used to displace the native interaction between bead-bound wild type protein. Mutations at the EloC pocket are highlighted in yellow, mutations at helix α5 interface are highlighted in purple and mutations at helix α2 interface are highlighted in green. The fitting and calculation of IC_{50} values was performed with GraphPad Prism 7 software and the error bars represent the error in the fitting function. **B)** Dose-response curves of the raw AlphaLISA data showing representative displacers from the three interaction areas in comparison with VBC wt. The experiments were performed in quadruplicate and the data points shown are the average values for each concentration. The error bars represent the standard deviation of each point.

2.2.5.3 ITC experiments

Based on the AlphaLISA results, protein mutants were selected for ITC experiments to determine K_d values. VBC mutants V181G, M105A, F109A, K159E and D187K were titrated, in order to quantify the importance of the hydrophobic environment in the EloC pocket and of the residues involved in the electrostatic interactions. There was also interest in quantifying the effect of mutating Leu3 and Pro5 in Cul2 by ITC, however, the mutant proteins were highly unstable resulting in unreliable data. The ITC experiments were performed titrating wt and mutant VBC proteins into Rbx1–Cul2 (Table 2.6 and

Appendix I). The same trend as in AlphaLISA was observed by ITC, with mutations M105A and F109A being the most disruptive and promoting, under these conditions, an increase of 20 and 37 fold in the K_d values, respectively, which was accompanied by a significant loss of binding enthalpy. On the other hand, the mutation K159E resulted in a 5-fold decrease of the affinity and the mutations of V181G and D187K decreased the binding affinity by about 3-fold.

Table 2.6 - Isothermal titration calorimetry data. K_d values and thermodynamic parameters for the interaction between VBC mutant proteins and Rbx1−Cul2 determined by ITC. VBC proteins (100 μM) were titrated into 10 μM of Rbx1–Cul2. ITC titrations performed in the exact same conditions are compared. VBC M105A and F109A titrations were performed at 303 K to obtain better quality data, as their interactions are of lower affinity, hence lower ΔH. All other titrations were performed at 298 K. The errors in the table reflect the quality of the fitting.

VBC	K_D (nM)	Temperature (K)	ΔG (cal/mol)	ΔH (cal/mol)	$T\Delta S$ (cal/mol)	N
wt	10 ± 3	298	-10958 ± 198	-12700 ± 177	-1742 ± 266	1
K159E	54 ± 9	298	-9920 ± 94	-11900 ± 174	-1930 ± 198	0.9
D187K	27 ± 9	298	-10339 ± 197	-11300 ± 264	-961 ± 330	0.8
V181G	28 ± 7	298	-10315 ± 154	-9868 ± 174	447 ± 232	0.9
wt	9 ± 1	303	-11017 ± 63	-14800 ± 61	-3783 ± 88	0.9
M105A	317 ± 39	303	-8873 ± 74	-9870 ± 141	-993 ± 159	0.9
F109A	177 ± 16	303	-9221 ± 53	-9260 ± 79	-41 ± 95	1

Together, the biophysical data are consistent with a ranking of hotspots at the Cul2–VBC PPI, being Met105 and Phe109, at the peripheral hydrophobic pocket of EloC, identified as the most critical amongst the mutations tested.

2.2.6 Swapping residues and selectivity for Cul2 *versus* Cul5

The mechanism behind the selectivity between recruitment of Cul2 *versus* Cul5 by EloBC-containing CRL complexes remains elusive, although

explored in several publications.[82,97,127] One of the existing hypotheses is that the electrostatic patch on pVHL surface is responsible for recruiting Cul2,[82] nonetheless, the data presented in the previous section shows that by inverting the charge of these two residues, and therefore, inverting the electrostatics patch charge, VBC still binds to Cul2. To deepen understanding of selectivity determinants, the $CRL5^{SOCS2}$ complex was studied in parallel. $CRL5^{SOCS2}$ is one of the Cul5 ligases in which SOCS2 is the substrate receptor and EloBC is the adaptor subunit (SBC).[116,128] The choice of this complex was motivated by the existence of a crystal structure of $Cul5_{NTD}$-SBC.[93] By ITC, it was observed that, despite using the same adaptor subunit, SBC does not bind Cul2. This intriguing observation led to the investigation of whether by mutating a combination of amino acid residues it would be possible to swap the selectivity of SBC with VBC for Cullin, or, in other words, rescue the binding of SBC to Cul2 (gain-of-function mutation) and, at the same time decrease the affinity to Cul5 (loss-of-function mutation) and vice-versa. The structural analysis of Cul2–VBC *versus* Cul5–SBC (Figure 2.12A) revealed that Arg186 in SOCS2 (Ser183 in pVHL) would clash with Gln111 in Cul2, which would explain the lack of interaction observed. In addition, it also showed that Lys159 and Asp187 in pVHL, important to the Cul2–VBC PPI, correspond to a glutamine (Gln164) and a tyrosine (Tyr190) in SOCS2, respectively. In other words, a positive and a negative amino acid residue in pVHL are replaced with two neutral amino acids in SOCS2, conferring a significantly different interface. Gln164 is involved in a hydrogen bond with Asn108 from EloC, while Tyr190 establishes an intra-molecular electrostatic interaction with Arg168. In addition, in SBC, Arg186 establishes hydrogen bonds with the backbone of Thr117 and with Gln113 of Cul5 (Figure 2.12A). These analyses and hypotheses were settled on the

assumption that the binding mode at the Cullin–adaptor interface is conserved, as observed in the crystal structures. Based on these observations, triple mutants for VBC and SBC were generated: mutations K159Q, S183R, D187Y were performed in pVHL (V^{QRY}BC); and the reverse mutations Q164K, R186S and Y190D were performed in SOCS2 (S^{KSD}BC).

2.2.6.1 Site-directed mutagenesis, protein expression and purification

The mutations were performed by site-directed mutagenesis (the primers can be consulted in Chapter 6) and confirmed by DNA sequencing. Expression and purification of V^{QRY}BC was performed as for the wild type protein, according to the protocol already explained. SBC and S^{KSD}BC were expressed and purified according to a protocol published previously,[116] involving affinity chromatography, ion-exchange chromatography and size exclusion chromatography. V^{QRY}BC and S^{KSD}BC were analysed by mass spectrometry to confirm the success of the mutations.

2.2.6.2 AlphaLISA and ITC experiments

Initially these variants were tested in the AlphaLISA competition assay (Figure 2.12B), and the results suggested a switch in the recruitment. A drastic weakening of the interaction between V^{QRY}BC and Cul2 was observed, and, pleasingly, it was also observed that S^{KSD}BC rescued, to some extent, the binding to Cul2, compared to SBC. Based on these encouraging results, ITC experiments were performed to evaluate the loss or gain in function upon triple mutation (Figure 2.12C-F and Table 2.7). A 25-fold loss in the binding affinity of S^{KSD}BC to Cul5 (K_d from 10 nM to 250 nM) was observed alongside with a partial rescue of the binding to Cul2 (K_d = 2.5 μM). Consistently, total loss of

binding between Cul2 and $V^{QRY}BC$ was observed, together with partial rescue of binding to Cul5 (K_d = 2.9 μM). The partially rescued non-native interactions confirmed a switch in the ability to recruit Cullin, albeit with K_d values that were weaker than the ones measured for the native interactions (Figure 2.12G). Interestingly, the loss of binding affinity of $S^{KSD}BC$ to Cul5 was accompanied by an increase in the enthalpy of the reaction, suggesting an interaction with a thermodynamic profile closer to that of Cul2–VBC (Figure 2.12C).

Taken together, these findings provide new insights into the selectivity of VBC and SBC for Cul2 and Cul5, respectively, identifying amino acid positions that direct this selectivity and mutations that enable switching Cullin recruitment within CRLs.

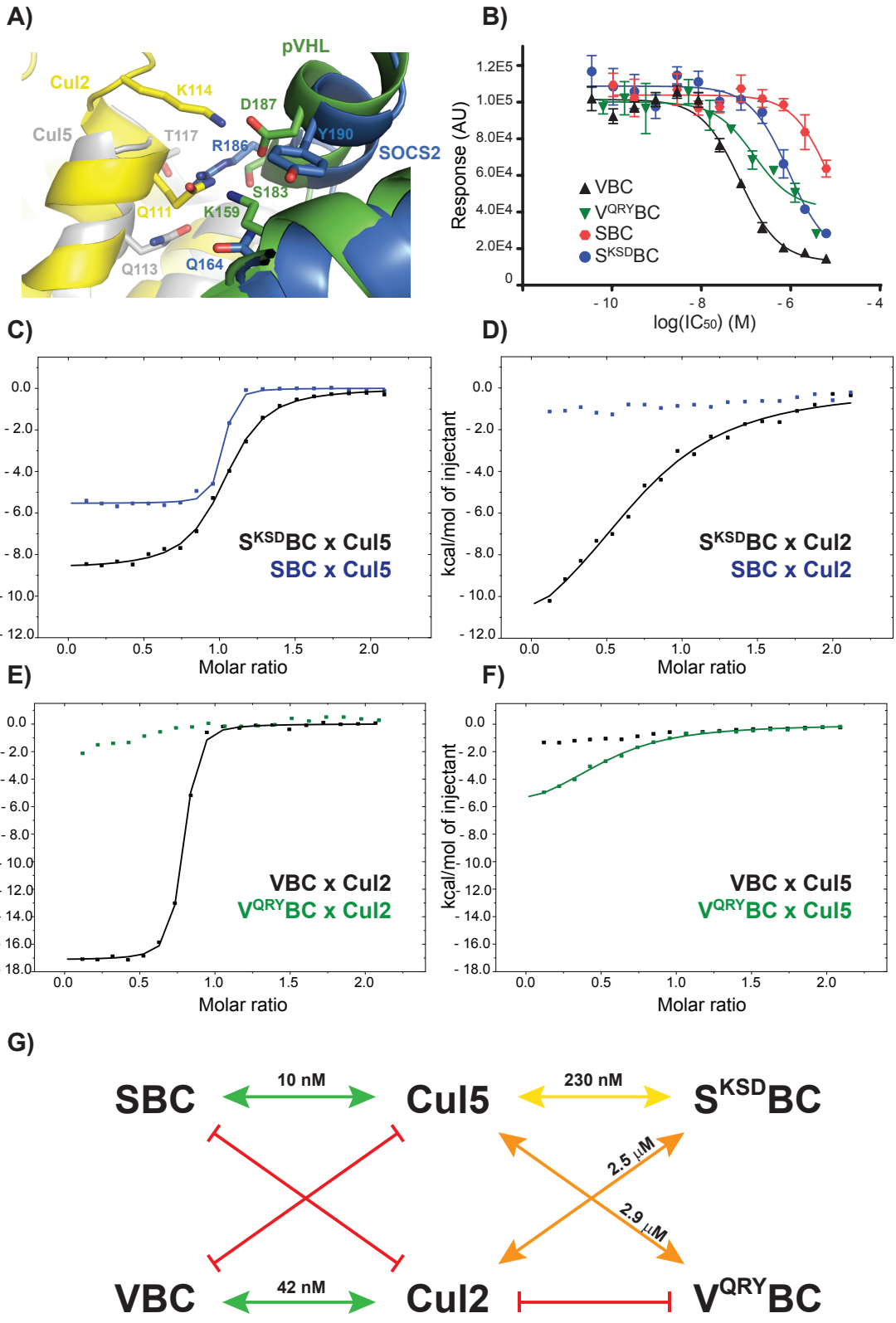

A)
pVHL
Cul2
K114
D187
Cul5
T117
R186
Y190
SOCS2
S183
Q111
K159
Q113
Q164
B)
Response (AU)
log(IC50) (M)
VBC
VQRYBC
SBC
SKSDBC
C)
kcal/mol of injectant
Molar ratio
SKSDBC x Cul5
SBC x Cul5
D)
SKSDBC x Cul2
SBC x Cul2
E)
VBC x Cul2
VQRYBC x Cul2
F)
VBC x Cul5
VQRYBC x Cul5
G)
SBC
10 nM
Cul5
230 nM
SKSDBC
2.5 μM
2.9 μM
VBC
42 nM
Cul2
VQRYBC

Figure 2.12 - Swapping residues and selectivity for Cul2 or Cul5. A) Structures of Cul5–SBC and Cul2–VBC aligned by the EloC subunit show residues involved in the electrostatic network between substrate receptor and Cullin, in both cases. Arg186 (SOCS2) clashes with Gln111 in Cul2 suggesting a potential reason for SBC not binding Cul2. **B)** AlphaLISA data shows loss of binding affinity of $V^{QRY}BC$ towards Cul2 and rescue of binding of $S^{KSD}BC$ towards Cul2. The experiments were performed in quadruplicate and the data shown are an averaged value. The error bars represent the standard deviation. The fitting was performed with GraphPad Prism 7 software. **C)-F)** Isothermal titration calorimetry data. Titrant solution (200 μM) was diluted into 20 μM of titrate over 19 injections of 2 μl at 303 K. **C)** Titration of $Cul5_{NTD}$ into SBC (blue) and $S^{KSD}BC$ (black). **D)** Titration of SBC (blue) and $S^{KSD}BC$ (black) into Cul2. **E)** Titration of $V^{QRY}BC$ (green) and VBC (black) into Cul2. **F)** Titration of $Cul5_{NTD}$ into $V^{QRY}BC$ (green) and VBC (black). **G)** Summary of the results obtained in the biophysical experiments

Table 2.7 - Isothermal Titration Calorimetry processed data of the experiments of the swapped domains. 200 μM of titrant diluted into 20 μM of titrate over 19 injections of 2 μl at 303 K. The errors refer to the fitting function.

	K_d (M)	Δ*G* (cal/mol)	Δ*H* (cal/mol)	TΔ*S* (cal/mol)
VBC x Cul2	$4.2 \times 10^{-8} \pm 5.0 \times 10^{-9}$	-10068 ± 70	-17210 ± 113	-7142 ± 133
$V^{QRY}BC$ x Cul2	ND	ND	ND	ND
SBC x Cul2	ND	ND	ND	ND
$S^{KSD}BC$ x Cul2	$2.5 \times 10^{-6} \pm 2.7 \times 10^{-7}$	-7652 ± 64	-12980 ± 524	-5328 ± 528
VBC x Cul5	ND	ND	ND	ND
$V^{QRY}BC$ x Cul5	$2.9 \times 10^{-6} \pm 3.0 \times 10^{-7}$	-7554 ± 59	-7400 ± 386	154 ± 39
SBC x Cul5	$9.8 \times 10^{-9} \pm 2.1 \times 10^{-9}$	-10936 ± 124	-5161 ± 33	5775 ± 128
$S^{KSD}BC$ x Cul5	$2.3 \times 10^{-7} \pm 3.2 \times 10^{-8}$	-9058 ± 81	-8354 ± 109	704 ± 136

2.3 Discussion

Ubiquitin ligases catalyse the transfer of ubiquitin to target proteins, which triggers a cellular response that varies according to the length and topology of the ubiquitin chain. Despite significant progress towards understanding the mechanisms underlying E3 ligase function, regulation and ubiquitination activity,[15] many aspects regarding the mechanistic details of specific interactions within the CRL complexes remain elusive. In fact, structural information at atomic level of the function and mechanism of the whole CRL machines remains limited.

The new structure of a full-length Cul2 complex provides insights that contribute to the understanding of the ubiquitination mechanism. It reveals flexibility of the central Cullin scaffold and corroborates the existence of hinge bending within the protein, which is deemed key for the ligase activity.[101] The hinge points found in the linkers between the Cullin repeats (first and second and second and third), allow the first and third helical bundles to come close or apart, alternating between an elongated and contracted shape. The other hinge point, located between the NTD and CTD, allows the rearrangement of these two domains and, in fact, it was observed that in the structure solved, the CTD of Cul2 is considerably moved towards the NTD, when compared to other CRL structures. Recent molecular dynamics studies have correlated these movements and proposed a relationship between Cullin and Rbx1 dynamics and the ubiquitin chain length and topology built on the substrate.[105] The new $CRL2^{VHL}$ structure presents for the first time the interface between the E2-recruiting RING, Rbx1, and Cul2, in which Rbx1 adopts a new orientation. Superposition with other Rbx1–Cullin structures suggests a pose in a trajectory

from a closed to an open form, which is postulated to capture a route from the inactive to the active conformation of the ligase.

CRLs must function as highly dynamic multi-subunit complexes. In fact, these complexes rely on their inherent flexibility to bring together substrate and ubiquitin that would otherwise be more than 50 Å apart, as found in some crystal structures.[84] Flexibility is also essential to accommodate multiple attachment of ubiquitin at distinct positions in the chain during catalytic elongation cycles. This feature is dictated by the relative organisation and dynamics of CRL subunits and their ability to work in concert. The new $CRL2^{VHL}$ structure, additionally to the intra- and the inter-subunit movements observed in comparison to other CRL structures, recapitulates many such dynamic features. For instance, the impossibility to obtain crystals in the absence of the HIF-1α 19-mer peptide suggests that the presence of the substrate was needed to confer some conformational arrangement that facilitates crystallisation. On the other hand, it is hypothesised that the substrate-binding domain of pVHL is highly flexible and might be present in more than one orientation in the crystal, explaining the lack of substantial electron density in this area. Although not captured by a single-structure crystallographic model, this observation suggests a functionally important dynamics of pVHL that could be driven by the presence of the bound substrate (HIF-1α peptide). Nussinov and others[129,130] have elucidated pVHL's inter-domain flexibility through the presence of a linker region containing a conserved proline residue, which was proposed to act as a hinge, allowing the rotation of the substrate binding domain to position the substrate accurately for ubiquitination. In addition to the above, the high average B-factors observed after refinement for the overall structure also corroborate these dynamic features. The resolution of the $CRL2^{VHL}$ structure limited the

amount of information that could be obtained, and, despite intensive efforts to improve the quality of the crystals and collect better quality datasets this was not possible.

In this chapter, the biophysical nature of a key PPI interface of CRL complexes was also explored. The interface established between Cul2 and the receptor and adaptor subunits is conserved and characterised by a large buried surface area. The structural features are reflected into characteristic thermodynamic signatures, including enthalpy-driven interaction, tight binding affinity and a large and negative change of heat capacity. It was determined that the interaction is contributed over extended regions, suggesting a spreading of hotspot residues across the interface. This work identified a number of hydrophobic residues in a distal pocket of EloC, which recognises the N-terminal tail of Cul2, as the major contributors to the PPI affinity. This information can guide the design of molecules to target this specific region within $CRL2^{VHL}$ and potentially other EloBC-containing CRL2 and CRL5 complexes. Moreover, the competitive AlphaLISA assay developed for these experiments can be of great value for the screening of modulators of the Cul2–VBC interaction.

The selective recruitment of Cullin in EloBC-containing complexes is still an intriguing feature due to the similarities of the two PPIs in Cul2 and Cul5 CRLs. Kamura *et al.* demonstrated the importance of the entire Cul-box region (about 20 amino acids) of the substrate recognition subunit in determining the selectivity for Cul2 vs. Cul5 recruitment.[97] In this work, upon swapping of three topologically conserved residues at the receptor-scaffold interface in VBC and SBC, the ability of these proteins to bind Cul5 and Cul2, respectively, was altered. This result constitutes an unprecedented proof-of-concept for switching

Cullin selectivity within CRLs by mutating individual amino acids and it can lead to interesting applications to elucidate CRLs biology and contribute to their validation as a drug target. It is likely that the Cullin selectivity is fine-tuned by other residues in addition to the ones identified in this work, considering they are not strictly conserved across all known Cul2 or Cul5 binders. In fact, it is known that mutation of Trp53 in Cul5 results in a 30-fold loss of affinity to SBC and this is due to an important interaction with Pro184 in SOCS2.[93] In Cul2, the corresponding residue to Trp53 is Ala48 and this could be one of the reasons why the binding affinity of $V^{QRY}BC$ to Cul5 is not fully rescued to the binding affinity of the native Cul5-SBC interaction.

In summary, the structural and biophysical data presented in this chapter shine new light into the structural assembly and dynamics functioning of CRL multi-subunit complexes and unveil mechanistic details and selectivity determinants for the Cul2-receptor protein-protein interaction.

CHAPTER 3

PROBING THE CRL2^{VHL} COMPLEX

3.1 Introduction

Peptides are attractive tools to target PPIs due to their ability to reproduce structural features of molecular recognition at usually large and flat surfaces. Encouraged by the potential of using peptides as chemical probes for PPIs,[131] the Cul2–VBC surface was explored with peptidic moieties designed to mimic the native interaction using sequences derived from Cul2.

Considering the practical aspects of exploring the Cul2–VBC interaction and developing peptides to target such interaction, it was deemed beneficial to have protocols that allow easy expression and purification of Cul2 and VBC with high yields to be used for biophysical and biochemical assays. VBC is routinely recombinantly expressed with high yields from bacteria and Rbx1–Cul2 can be expressed and purified from insect cells. However, expression in insect cells is a lengthy and costly procedure that yields relatively low amount of protein. The bacterial expression system, on the other hand, is cost-effective and fast, easy to handle and for most recombinant proteins it generates abundant yields.[132]

In this chapter, it is described work aimed at establishing procedures for Cul2 purification using a bacterial host and to develop peptides by structure-guided design to probe the Cul2–VBC interaction. The peptides were biophysically evaluated for binding towards VBC by a palette of biophysical techniques including BLI, DSF, ITC and NMR. The results pointed out important insights to further develop these peptides into higher-affinity ligands that could be used as chemical tools to deepen our understanding of the assembly and biological functioning of CRL complexes.

3.2 Results

3.2.1 Design, expression and optimisation efforts to obtain soluble and monomeric $Cul2_{NTD}$ constructs

The work focused on two truncations of Cul2 with different lengths, one comprising the first Cullin repeat (residues 1-163) and the other one comprising the full NTD (residues 1-384). The design of the constructs was based on the work of Nguyen *et al*.[82] However, since the goal was to study the interaction between the N-terminus of Cul2 and VBC, it was decided to have the solubilising tag at the C-terminus of the protein rather than at the N-terminus. Two solubilising tags, SUMO and MBP, were tested and a hexahistidine tag was added at the C-terminus (downstream of the solubilising tag) for purification purposes. Details of the cloning can be found in Chapter 6. Protein expression trials in *E. coli* were carried out for all constructs containing the different tags (Table 3.1). Parameters such as temperature and concentration of isopropyl-β-D-thiogalactopyranoside (IPTG) used to induce protein expression can influence protein expression in *E. coli* and can be tuned to achieve optimal expression conditions.[133] In this case, a range of temperatures, 16-25 °C, and concentrations of IPTG, 0.05-0.5 mM, were tested. It was observed that, although the protein was overexpressed in all cases, none of the constructs was soluble under any of the conditions tested. The next step was to attempt protein expression in the presence of GroEL and GroES chaperones (Chaperone plasmid set, TaKara). Molecular chaperones are known to be involved in protein folding *in vivo*,[134] therefore, by co-expressing a protein target in the presence of chaperones the recovery of expressed proteins in the soluble fraction is usually increased. To do this, the Cul2-expression vectors were transformed into BL21 *E. coli* cells containing pGro7 plasmid (see Chapter 6,

section 6.1.5). In this system, the expression of GroEL and GroES chaperones is regulated by arabinose and the concentration used was tuned for each of the systems (0.5-1.5 mg/ml). The presence of the chaperones enhanced the solubility of the MBP-tagged constructs and up to 50% of the protein was found in the soluble fraction. Conversely, despite being overexpressed, the SUMO-tagged constructs remained totally insoluble. Based on these results, the MBP-tagged constructs expressed in the presence of chaperones were chosen for further studies.

Table 3.1 – Summary of the Cul2 constructs tested and solubility results.

Construct	Tag	Chaperones	Solubility
$Cul2_{1-163}$	SUMO-6xHis	No	Insoluble
		Yes	Insoluble
	MBP-6xHis	No	Insoluble
		Yes	Soluble
$Cul2_{1-384}$	SUMO-6xHis	No	Insoluble
		Yes	Insoluble
	MBP-6xHis	No	Insoluble
		Yes	Soluble

The expression and purification efforts of MBP-tagged constructs of $Cul2_{1-163}$ and $Cul2_{1-384}$ will be described in the next sections.

3.2.1.1 Expression and purification of Cullin-2_{1-163}-MBP-6xHis

The Cullin-2_{1-163}-MBP-6xHis construct comprised the first cullin repeat, constituted by the N-terminal loop and the first five α-helices. After expression trials to determine the optimal conditions, Cul2 expression was induced with 0.5 mM IPTG and the cells were incubated overnight at 16 °C, 200 rpm shaking. The cells were harvested by centrifugation and lysed by French press prior to purification. Originally, the first step of purification consisted of an affinity

chromatography with a His trap (column packed with nickel sepharose resin). However, after several attempts at purification, it was observed that the amount of Cul2 recovered by gradient elution with imidazole was much lower than the amount present in the lysate. This was very intriguing, and after some literature search it was found that Cul2 binds nickel.[135] This explained the poor recovery of protein from the His trap column in the elution step at the expected imidazole concentration (~100 mM). Considering that the protein contained a maltose binding protein (MBP) as tag and the existence of MBP trap affinity purification columns (packed with dextrin sepharose resin), the His trap step was replaced with MBP trap affinity purification.

Protein expression, cell harvesting and cell lysis were performed as before and, this time, $Cul2_{1\text{-}163}$-MBP-6xHis was recovered by elution with maltose from the MBP trap column (Figure 3.1A). Next, a gel filtration step was performed and the size-exclusion chromatography (SEC) trace consisted of a single peak corresponding to Cul2 that, however, was eluted before the expected volume (Figure 3.1B). According to the molecular weight (MW) of the protein (63 kDa) and to the calibration of the column (performed using the gel filtration standard calibration kit, BioRad), monomeric $Cul2_{1\text{-}163}$-MBP-6xHis was expected to have an elution volume of ~70 ml in the HiLoad 16/600 Superdex 200 column. Instead, $Cul2_{1\text{-}163}$-MBP-6xHis eluted at 50 ml, which corresponds to a MW of approximately 440 kDa, suggesting that the protein was purified as an aggregate. In the chromatogram it was also visible a second peak that eluted at the volume corresponding to a species of about 45 kDa. It was speculated that this could be the result of spontaneous cleavage of MBP (as its MW is 45 kDa). To prevent Cul2 aggregation, given that Cul2 binds Ni^{2+}, one idea was to dilute the aggregate sample in buffer containing 10 mM $NiCl_2$ as

this could stabilise the monomeric species. After dilution, the sample was analysed again by gel filtration but, unfortunately, no monomeric species was observed also in the presence of $NiCl_2$.

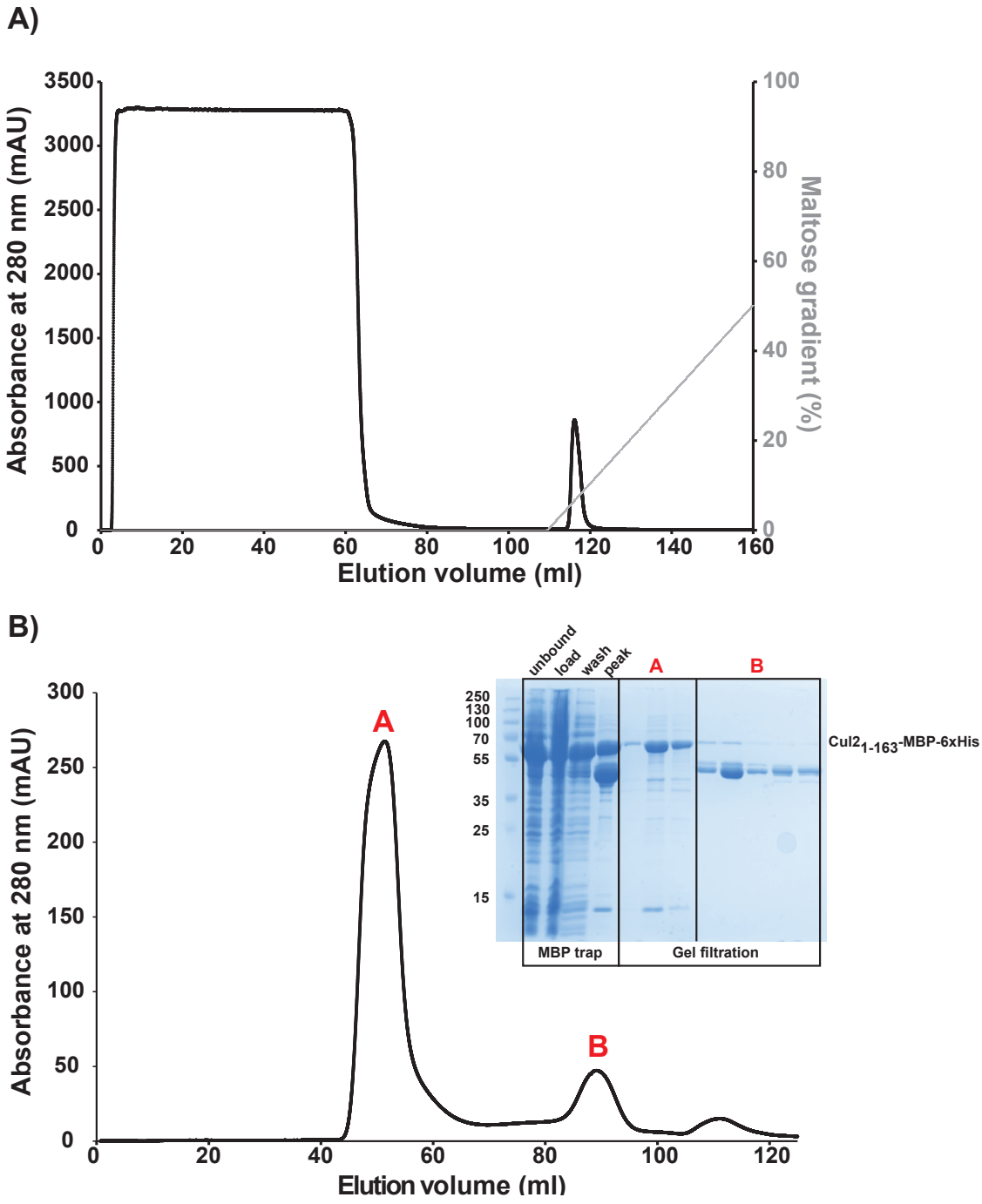
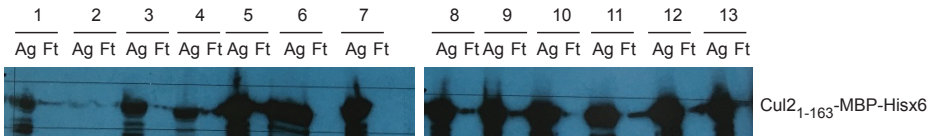


Figure 3.1 – Chromatograms of the purification of Cul2$_{1-163}$-MBP-6xHis and respective SDS-PAGE. A) The protein is eluted from the MBP trap with a gradient of maltose (0-5 mM). **B)** Size-exclusion chromatography profile of Cul2$_{1-163}$-MBP-6xHis after the MBP trap. The SDS-PAGE confirms the identity of Cul2$_{1-163}$-MBP-6xHis (A) based on the MW and the presence of a species with MW ~45 kDa (B).

In order to identify conditions that could stabilise the monomeric species, a solubility assay was performed, based on the protocol described by Bondos *et al*.[136] The assay involved diluting the aggregate in buffers containing different additives. Then, the protein was spun down in centricons of 100 kDa cutoff, the idea being that the aggregate would be retained (expected MW according to SEC > 440 kDa), whereas monomeric protein (MW 63 kDa) should go through the membrane. Samples from the retained phase and from the filtrate were loaded on a SDS-PAGE and Cul2 was identified by immunoblot. In this way, conditions that led to a stabilisation of the monomeric species could be identified. The results showed that the greatest amount of Cul2 in the flow-through was obtained in the presence of the detergent BRIJ58, glucose or KCl (conditions 5, 9 and 11, Figure 3.2), hinting at the presence of monomeric species. According to that, small-scale purifications were performed in the presence of each of those additives. In SEC, the monomeric species was only observed in the presence of BRIJ58. However only approximately 25% of the total protein preparation was found in a monomeric state and it was observed by SEC that upon isolation of the monomeric species, it would aggregate again. In conclusion, none of the additives tested allowed durable stabilisation of monomeric $Cul2_{1\text{-}163}$-MBP-6xHis.



Condition	Buffer	Additive
1	50 mM HEPES pH 7.0, 100 mM NaCl, 1 mM TCEP	-
2		40% ethanol
3		10% DMSO
4		5% PEG
5		0.5% BRIJ58
6		0.5 M L-arginine
7		0.5 M proline
8		1% glycine
9		1 M glucose
10		0.1 M $CaCl_2$
11		1 M KCl
12		0.15 M $(NH_4)_2SO_4$
13		1 M urea

Figure 3.2 – Solubilisation assay. Different additives were tested to favour the monomeric fraction of Cul2$_{1-163}$-MBP-6xHis. Samples from the aggregate (Ag) and flow through (Ft) were analysed by immunoblot with anti-Cullin-2 antibody at 1:2000.

Another avenue explored was the denaturation of purified aggregate of Cul2 in the presence of 8 M urea, followed by refolding using different conditions (L-arginine, L-proline, glucose and ammonium sulphate) and methods (dilution, refolding on the column and dialysis). The refolding was assessed by SEC. For example, when the refolding was attempted on the column (Figure 3.3), contrarily to the aggregate, the protein was eluted closer to the predicted volume according to its molecular weight. In the Superdex 200 Increase column (analytical column), the expected elution volume for Cul2$_{1-163}$-MBP-6xHis is around 12 ml, which was approximately the elution volume observed for the samples tested, suggesting the protein had been refolded. To assess whether the refolded samples of Cul2 were active, they were tested for binding to VBC by co-elution gel filtration. It was expected that upon mixing, the two proteins would form a single complex and elute in the format of a single peak, however this did not happen, indicating that the protein was not active (data not shown).

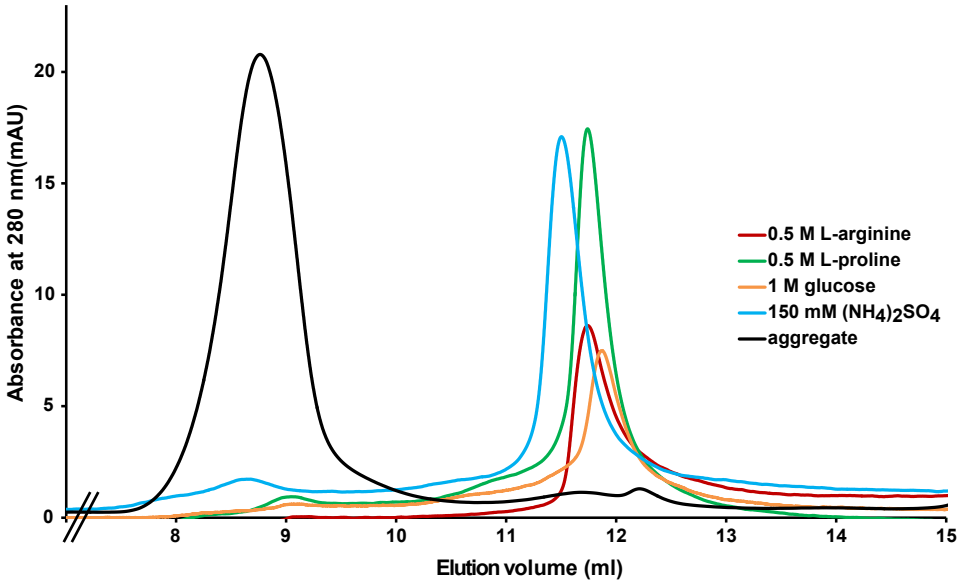


Figure 3.3 – Refolding experiments with Cul2. The aggregated sample was purified, denatured in 8 M urea and then refolded on the column (Superdex 200 Increase) with buffer (100 mM HEPES pH 7.0, 100 mM NaCl, 1 mM TCEP) containing different additives.

Despite the different additives and methods of refolding tested, no active monomeric Cul2 was obtained.

3.2.1.2 Expression and purification of Cullin-2$_{1-384}$-MBP-6xHis

Considering the previous results, the idea of a longer construct was contemplated as a way of blocking an exposed hydrophobic surface at the end of the first helical bundle, which is usually packed against another helical bundle. Expression of Cul2$_{1-384}$-MBP-6xHis comprising the full NTD in the presence of the chaperones was induced with 0.5 mM IPTG and carried out at 16 °C overnight. Cells were harvested by centrifugation, lysed by French press and loaded on the MBP trap. The peak collected from elution with maltose was loaded on a gel filtration and the SEC trace revealed the presence of multiple species of Cul2$_{1-384}$-MBP-6xHis. The most abundant species was eluted in the void volume, suggesting also this construct was prone to aggregation

(Figure 3.4). The addition of 5% glycerol in the purification buffers was tested to prevent aggregation but no improvement was observed. Moreover, as with the shorter construct, it was observed that the aggregation was equilibrium-dependent *i.e.* upon isolation of the monomeric species the sample would start aggregating.

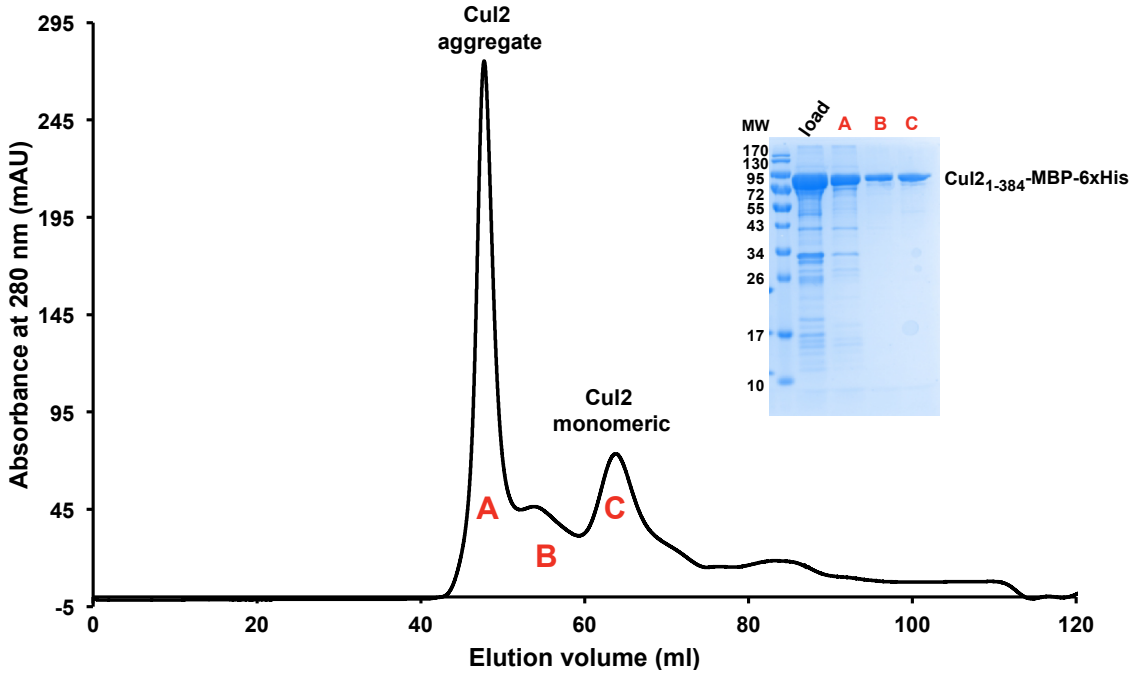


Figure 3.4 – Size exclusion chromatography trace of Cul2$_{1-384}$-MBP-6xHis. The chromatogram and the SDS-PAGE suggest the presence of multiple species of Cul2$_{1-384}$-MBP-6xHis.

Previous work by Li *et al.*[137] reported the mutation of residues Val367 and Leu371, two hydrophobic solvent-exposed residues, to arginine and aspartate, respectively, to improve the solubility of Cul1. The same principle was applied to this construct of Cul2, hoping to revert the aggregation issue. To decide which residues to mutate, Cul1 and Cul2 were structurally aligned and residues Val340 and Leu344 of Cul2 were identified as corresponding to residues Val 367 and Leu371 in Cul1. In addition to Val340 and Leu344, it was decided to also mutate Val384. The valine residues were mutated to arginine

and the leucine was mutated to aspartate ($Cul2_{1-384}^{RDR}$). The mutations were performed by site-directed mutagenesis and confirmed by DNA sequencing.

The protein was expressed using the same protocol as above. $Cul2_{1-384}^{RDR}$-MBP-6xHis was eluted from the MBP trap column as a very sharp peak. When loaded on the gel filtration column (HiLoad 16/600 Superdex 200), the chromatographic profile suggested that the sample was aggregated as the peak eluted shortly after the void volume, therefore, implying that the mutations did not improve the aggregation problem.

Another condition tested for the purification of $Cul2_{1-384}^{RDR}$-MBP-6xHis was to lower the NaCl concentration to 100 mM in all the buffers, an attempt to disfavour hydrophobic contacts and minimise aggregation. Surprisingly, upon loading of the MBP trap, part of the sample was eluted while washing the column, before starting the maltose gradient elution (Figure 3.5A). SDS-PAGE confirmed that this sample was Cullin-2_{1-384}^{RDR}-MBP-6xHis (sample A) and it was kept separated from the sample eluted with maltose (sample B). Sample A was run in the gel filtration column and, it was observed that approximately 40% of the sample was eluted at the expected volume, according to the respective molecular weight, suggesting the presence of monomeric species. On the contrary, in sample B only about 10% of sample was monomeric and the remaining ~90% eluted in the void volume. Thus, lowering the NaCl concentration did not help to enhance the population of monomeric fraction. However, somehow a fraction of monomeric Cul2 was obtained from this experiment and was used to try to understand what could help its stabilisation.

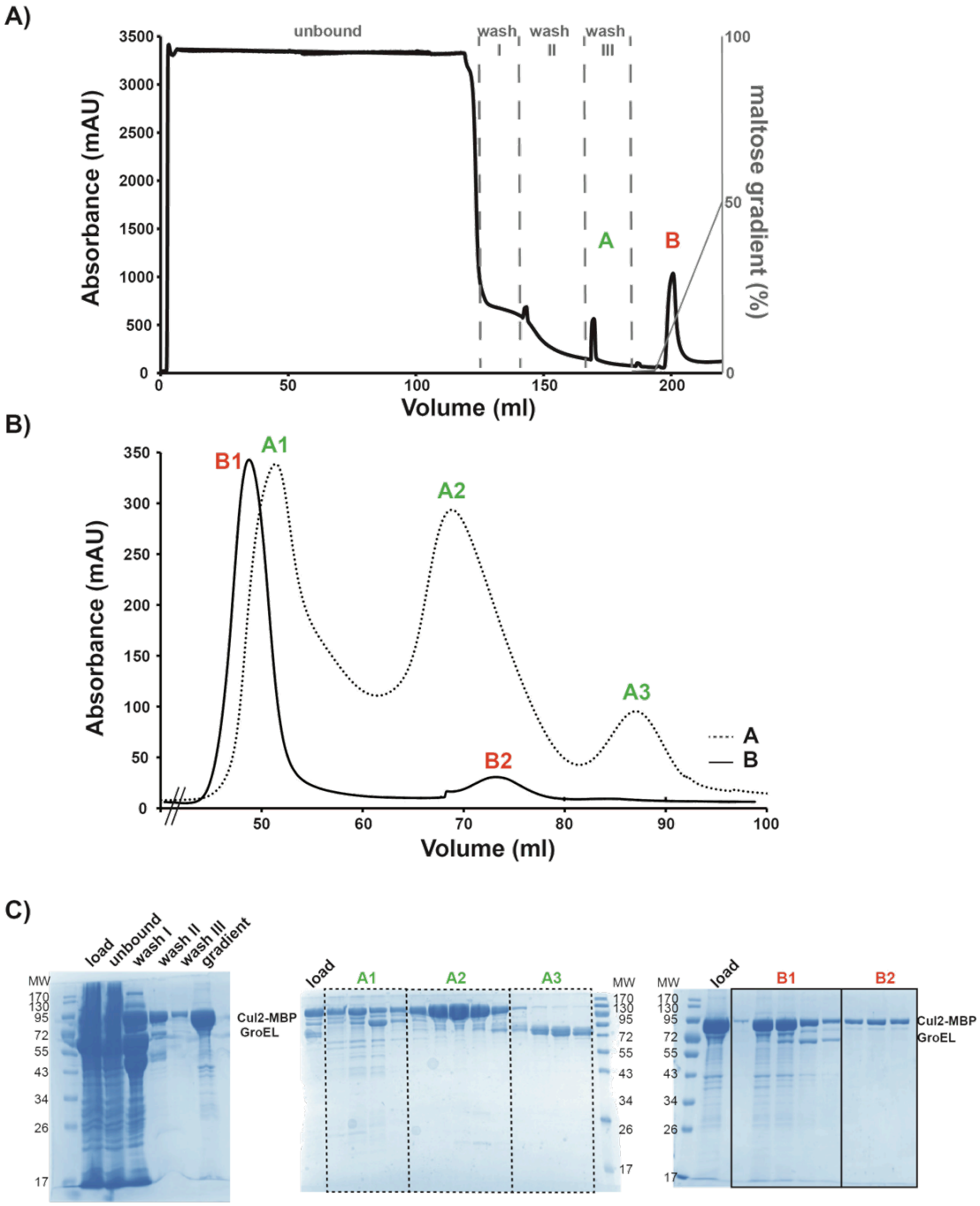


Figure 3.5 – Purification of Cul2$_{1\text{-}384}^{RDR}$-MBP-6xHis. A) MBP trap elution profile, where, unexpectedly, part of the sample was eluted during the wash step (sample A). **B)** Size exclusion chromatography of samples A and B. Sample A contained a great amount of monomeric Cul2$_{1\text{-}384}$-MBP-6xHis (A2), while most of sample B was aggregate. **C)** SDS-PAGE from samples from MBP trap and gel filtration runs confirms the identity of the peaks as being Cul2$_{1\text{-}384}^{RDR}$-MBP-6xHis.

To assess whether the monomeric species of Cullin-2$_{1\text{-}384}^{RDR}$-MBP-6xHis from sample A was active *i.e.* binding to VBC, a co-elution gel filtration experiment was run. Cullin-2$_{1\text{-}384}^{RDR}$-MBP-6xHis and VBC were mixed in a 1:1

ratio and the elution profile was analysed. Comparing to the individual sample's profile there was a new peak eluting before the peak of Cul2, consistent with a higher molecular weight sample, as would be expected for Cul2–VBC (Figure 3.6). The SDS-PAGE confirmed the presence of VBC under this first peak (which did not happen when VBC was loaded alone), corroborating the assembly of the Cul2–VBC complex. However, based on the chromatogram, it was estimated that only about 10% of VBC was forming a complex with Cul2. This was consistent with aggregation, which was also suggested by the presence of some higher molecular weight species in the chromatographic trace of Cul2 (Figure 3.6, yellow trace) and of the mixture (Figure 3.6, green trace).

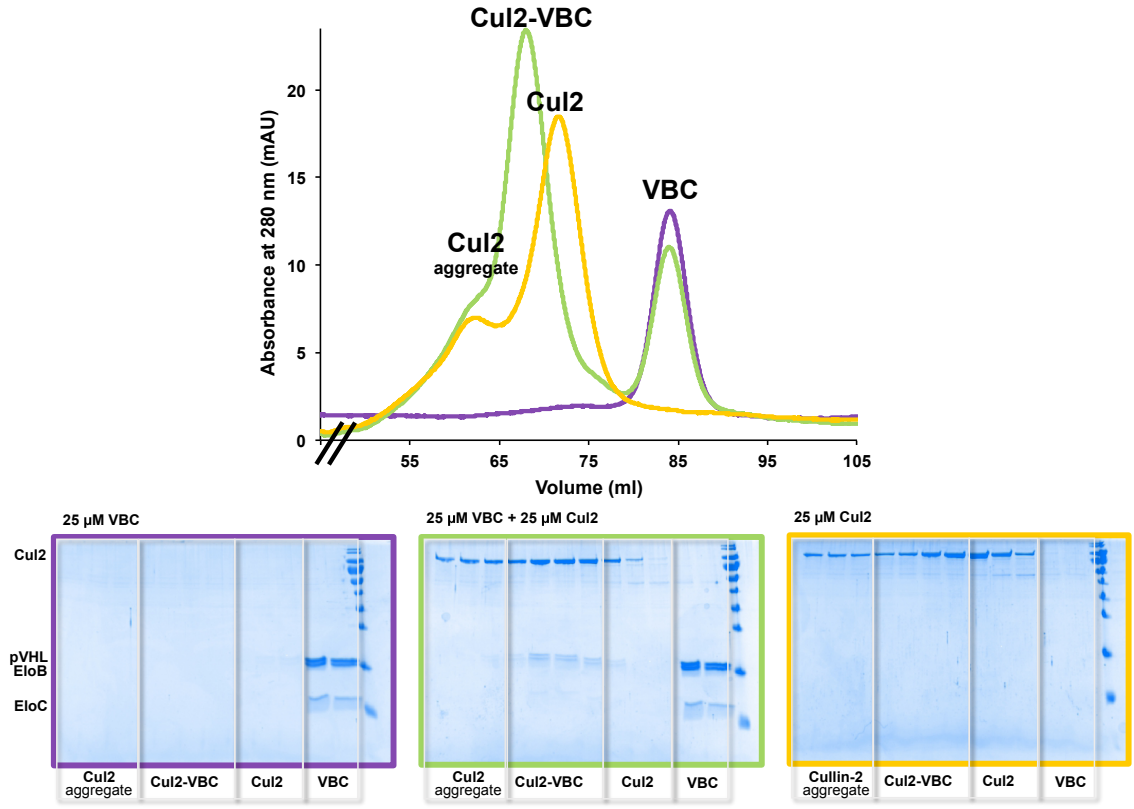


Figure 3.6 – Co-elution gel filtration of Cul2 and VBC. Monomeric $Cul2_{1\text{-}384}^{RDR}$-MBP-6xHis was mixed with VBC in a 1:1 ratio and the size-exclusion chromatography trace (green) was analysed and compared to the chromatographic traces of VBC (purple) and monomeric $Cul2_{1\text{-}384}^{RDR}$-MBP-6xHis (yellow).

Apart from the co-elution gel filtration, BLI and ITC experiments were carried out with the monomeric fraction of Cul2 isolated from this experiment. However, no clear binding could be detected and this might be explained by the high content of Cul2 aggregate present in the samples used, as it was seen in the gel filtration traces.

In an effort to stabilise monomeric Cul2, a thermal shift screening of different buffer conditions was performed. The more stable the protein is, the higher the T_M is expected.

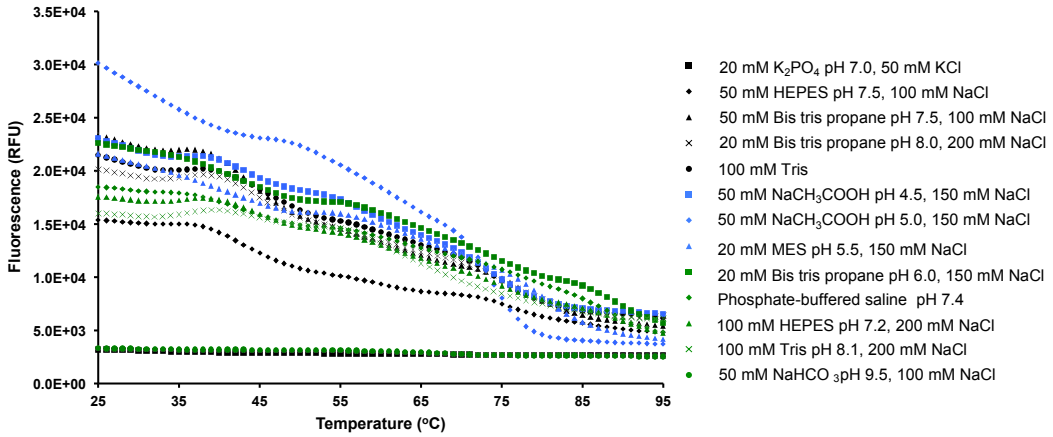


Figure 3.7 – Thermal shift buffer screening. Amplification curves of 5 μM $Cul2_{1\text{-}384}^{RDR}$-MBP-6xHis in the presence of different additives.

The results of the buffer screening were inconclusive as no clear melting curve was detected in any of the conditions tested (Figure 3.7). It was observed that the fluorescence signal at the beginning of the experiment was already extremely high. The reasoning for this observation was that Cul2 is a very hydrophobic protein and, consequently, the typical melting curves consisting of a peak of fluorescence intensity upon denaturation of the protein were not observed and no conclusion in regards to the thermal stability of the protein could be made.

The purification in the conditions that led to the isolation of a considerable amount of monomeric Cul2 was repeated several times but it was not possible to reproduce the original result.

Other parameters like lysis and purification conditions were tested with the $Cul2_{1\text{-}384}^{RDR}$-MBP-6xHis construct. Different buffers, varying the salt concentration and including additives such as Tween-20 and $NiCl_2$ were tested but there was no increase of the monomeric fraction according to the SEC profile. In addition, as a great amount of soluble aggregate of Cul2 could be purified and assuming the aggregation was reversible, the aggregate was diluted and additives such as Arg•HCl, $CoCl_2$ and glycerol were tested but no improvement was observed either. The solubility assays performed for the shorter construct were also performed in this case, however, it was observed that in none of the conditions the protein went through the membrane, suggesting it was always present as an aggregate.

Lastly, the denaturation and refolding protocols were also tested for Cullin-$2_{1\text{-}384}^{RDR}$-MBP-6xHis. In this case, the refolding was only attempted on the column and by dialysis, using arginine as additive. Nonetheless, it was observed that it was not possible to refold the longer construct or, if refolded, it was immediately aggregating.

In summary, despite the intensive efforts, unfortunately it was not possible to achieve a soluble, monomeric and stable construct of Cul2 expressed in *E. coli*.

3.2.2 3- to 11-mer Cul2 peptides

Considering the important role of the N-terminal loop of Cul2 in establishing the interaction with VBC, and given the inability to express soluble $Cul2_{NTD}$ alone, it was contemplated that peptides able to reproduce the N-terminal tail could recapitulate the interaction, providing tools to develop chemical probes and target this PPI. Based on the structural analysis of this interface, peptides varying from three to eleven residues were synthesised and tested for binding to VBC.

3.2.2.1 Peptide synthesis and purification

The peptides were synthesised in an INTAVIS automated peptide synthesiser using solid phase Fmoc chemistry. The synthesis consisted of several cycles of deprotection, washing and coupling steps; each cycle resulted in the incorporation of one amino acid. After the final deprotection step, the peptides were cleaved from the resin using a solution of TFA, water and triisopropylsilane (TIS) (92.5:2.5:5). The peptides were obtained as C-terminal amides and were purified by HPLC in basic conditions (0.1% NH_4OH) in a gradient of 0-100% of acetonitrile in water over 15 minutes. The purity and identity of the peptides was determined by analytical HPLC-MS (see Chapter 6).

3.2.2.2 Biolayer interferometry

The nine peptides synthesised were tested for binding to VBC by BLI at increasing concentrations, from 27 μM to 7 mM. For this experiment, biotinylated VBC (see Chapter 6, section 6.1.14) was attached to streptavidin-coated tips. The 9-mer and 11-mer peptides could not be tested

because they were insoluble in the conditions of the experiment. For the remaining peptides the results showed that at least six amino acids were needed in order to observe binding towards VBC (Figure 3.8 and Table 3.2). The binding was recapitulated with the 6-, 7-, 8- and 10-mer peptides and the 8-mer peptide showed the best response in terms of fitting and reproducibility. The same experiment was performed in the presence of VBC in complex with the HIF-1α 19-mer peptide (VBCH) and the results obtained were similar (Table 3.2). Nevertheless, in both cases the binding affinities determined were in the millimolar range, therefore, the next step was to enhance the binding affinity of the peptides towards VBC.

Table 3.2 – Binding affinities of N-terminal Cul2 peptides to VBC and VBCH determined by BLI.

Peptide	K_d VBC (mM)	K_D VBCH (mM)
MSL	no binding	no binding
MSLK	no binding	no binding
MSLKP	no binding	no binding
MSLKPR	3.3 ± 0.3	11.0 ± 0.9
MSLKPRV	3.5 ± 0.2	5.3 ± 0.4
MSLKPRVV	2.0 ± 0.2	4.5 ± 0.5
MSLKPRVVD	ND	ND
MSLKPRVVDF	4.0 ± 0.3	5.8 ± 0.5
MSLKPRVVDFD	ND	ND

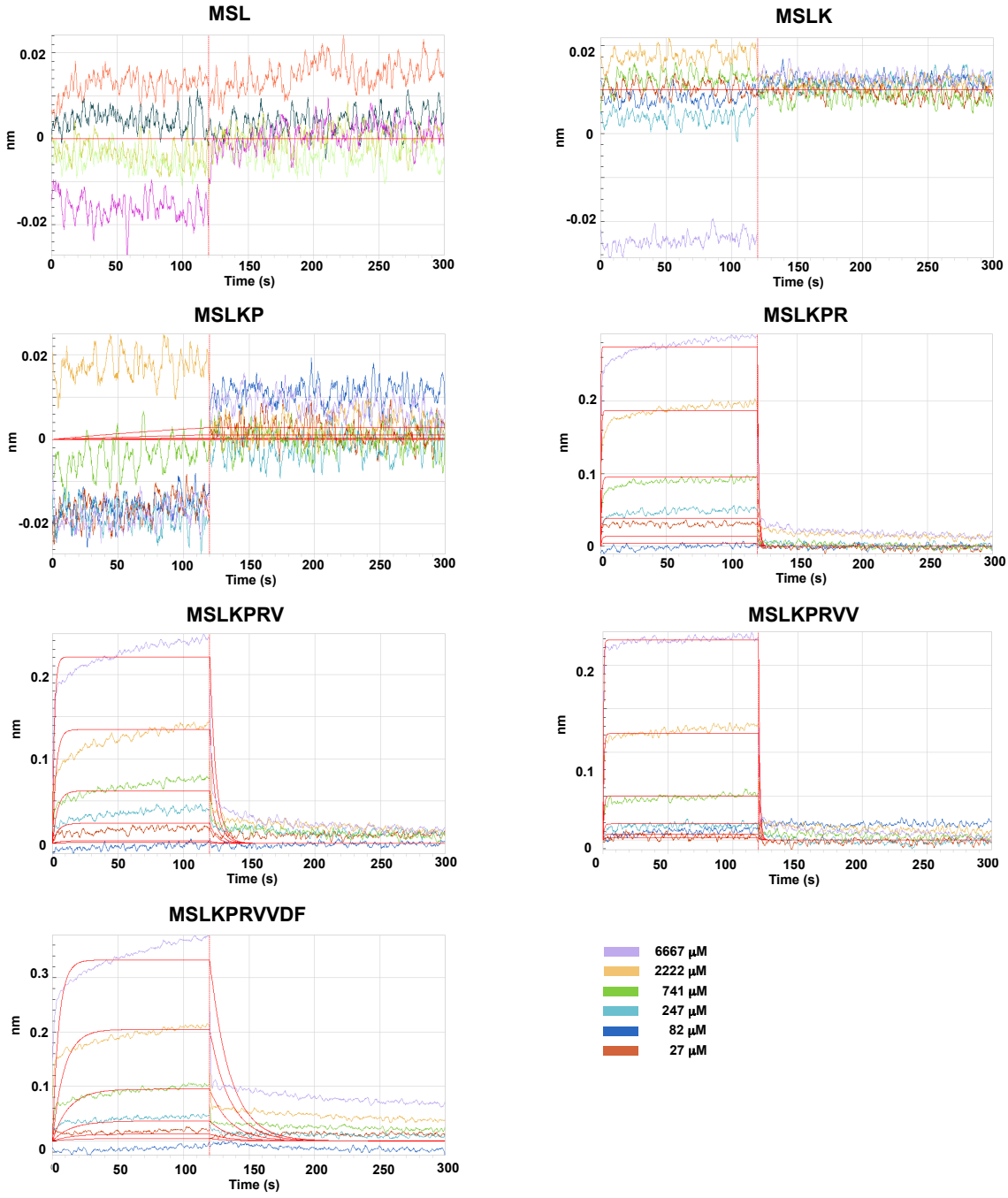


Figure 3.8 – Biolayer interferometry experiments with the N-terminal Cul2 peptides. Biotinylated VBC was immobilised on the tips and the peptides were tested in a concentration range from 27 µM to 7 mM. The different colours represent the different concentrations and corresponding responses. The assay was performed in 20 mM HEPES pH 7.6, 100 mM NaCl, 1 mM DTT, 0.02% Tween-20 at 25 °C.

3.2.3 8-mer peptide derivatives

Considering the 8-mer peptide (MSLKPRVV) had the best fitting (see Figure 3.8), tighter binding affinity and reproducibility in the previous BLI assay, it was chosen as template for a series of modifications to determine the requirements for the interaction and potentially increase the strength of binding.

An alanine scanning was performed to identify hotspots in the peptide. Furthermore, as the binding affinity of the 8-mer peptide was relatively weak (2 mM), a structure-based approach to try to enhance the affinity was employed. Considering the key role of the side chain of Leu3 from Cul2 accommodated in the EloC hydrophobic pocket, the replacement of leucine with bulkier side chain groups was tested. Thus, a small library of 8-mer peptides in which the leucine residue was substituted with natural and non-natural amino acids was designed and tested for binding to VBC. A total of 19 peptides were synthesised and purified as described in section 3.2.2.1. All the peptides were characterised by analytical HPLC-MS and the traces can be consulted in Chapter 6.

3.2.3.1 Alanine scanning

The alanine scanning consisted of replacing, one at a time, all the amino acid residues in the 8-mer peptide by an alanine residue. The peptides were then tested for binding to VBC by BLI, as described in section 3.2.2.2, and the sensorgrams were compared to the one of the parental peptide, considering the response signal (Figure 3.9). Based on the results, mutations that led to complete disruption of the binding event were identified. It was observed that by replacing Pro5 or Arg6 with alanine, the binding of the peptide towards VBC

was totally lost. This was in agreement with the structural analysis of Cul2–VBC discussed in the previous chapter.

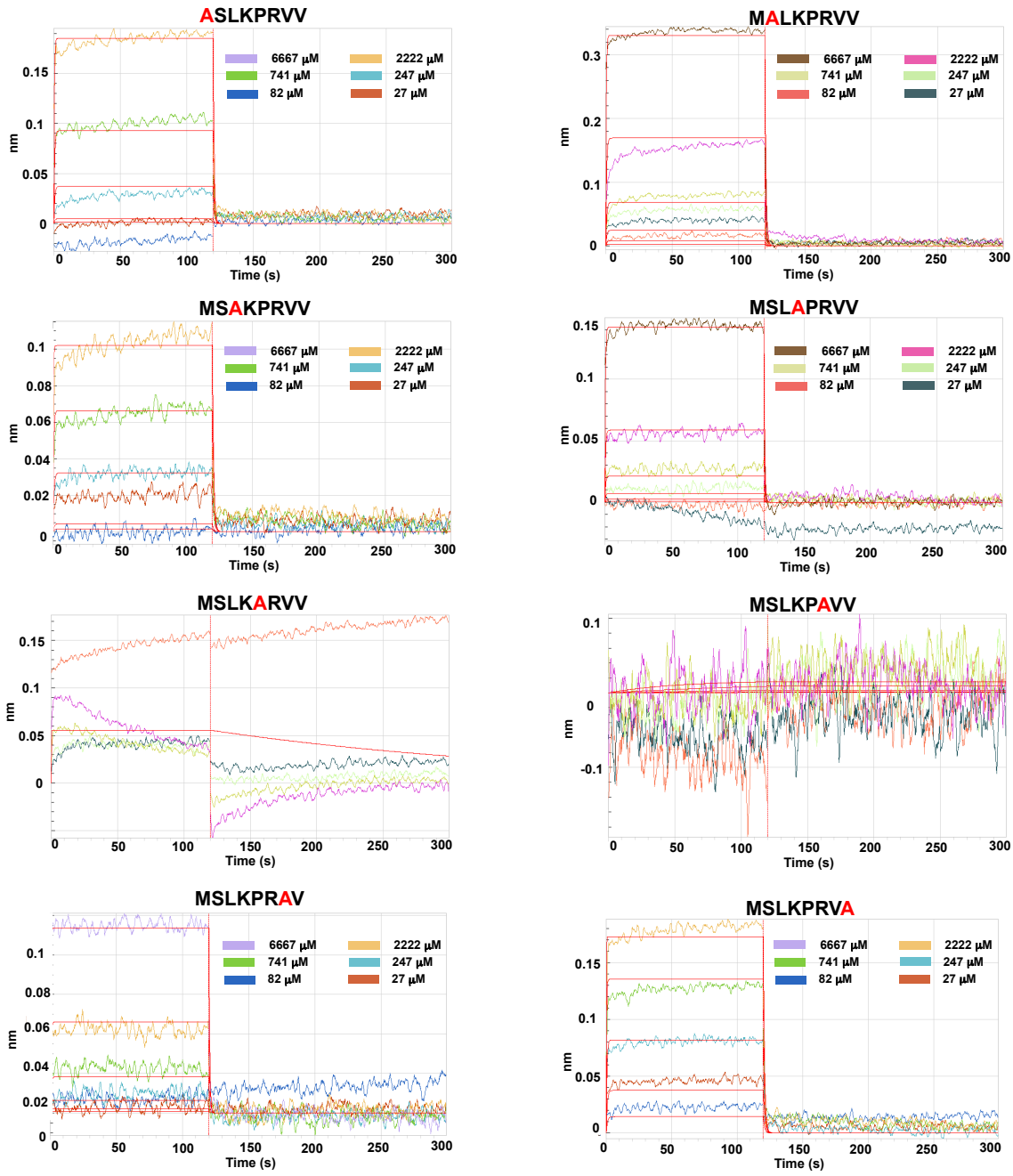


Figure 3.9 – Biolayer interferomery experiments with the alanine scan peptides. Biotinylated VBC was immobilised on the tips and the peptides were tested in a concentration range from 27 µM to 7 mM. The different colours represent the different concentrations and corresponding responses. The assay was performed in 20 mM HEPES pH 7.6, 100 mM NaCl, 1 mM DTT, 0.02% Tween-20 at 25 °C.

Indeed, Pro5 is thought to be responsible for the folding of the N-terminal loop of Cul2 upon itself and Arg6 responsible for keeping this conformational

arrangement by establishing intramolecular hydrogen bonds with the carbonyl groups from the backbone of Ser2 and Lys4. In addition, replacing Lys4 with alanine promoted 10-fold worsening of the binding affinity, however, the replacement of other amino acids in the peptide with alanine did not seem to perturb the interaction as much. This was an interesting result, suggesting that the structural pre-arrangement of the peptide was important to establish the interaction.

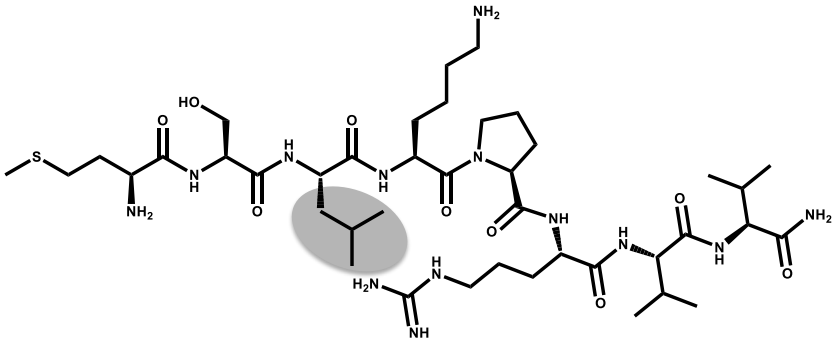
3.2.3.2 Leucine replacement

A total of 11 amino acids were used to replace Leu3 (Table 3.3). The main criteria for choosing these amino acids was to modulate the bulkiness of the side chain group, for example, leucine was replaced with phenylalanine and *t*-butylglycine, among others, to increase the volume occupied by the side chain. The only exception was arginine, which was chosen to try to catch an electrostatic contact at the bottom of the EloC pocket. As before, the peptides were tested by BLI for binding to VBC and VBCH.

The BLI results showed that the EloC pocket could accommodate all the derivative peptides, except when leucine was replaced by tryptophan, which was probably overly bulky and was found to decrease the binding affinity in comparison to the parental peptide (Table 3.3). Replacement of leucine with arginine, contrarily to what was predicted, did not improve the binding affinity. In addition, it was also observed that the replacement of leucine with *t*-butylglycine and dimethylcysteine (peptides **G** and **J**) led to an improvement of the binding affinity to the micromolar range. Particularly for peptide **J**, the K_d values determined were 0.3 mM and 0.5 mM for VBC and VBCH, respectively, which

encouraged further in-depth characterisation, with the aim to better understand its binding interaction.

Table 3.3 – Results from the BLI experiment with the peptide variants replacing leucine. The K_d of peptide E to VBC could not be determined (ND) as the data could not be fitted by the 1:1 model function, suggesting, perhaps, aggregation. Peptide **L** was only tested against VBC.



Peptide	Leu3 replacement	K_d (mM) VBC	K_d (mM) VBCH
A	Leucine	2.0 ± 0.2	4.5 ± 0.5
B	Tryptophan	18 ± 1	34 ± 5
C	Phenylalanine	3.1 ± 0.2	3.1 ± 0.4
D	Tyrosine	7.4 ± 0.4	8.9 ± 0.2
E	Norleucine	ND	1.7 ± 0.2
F	Norvaline	1.9 ± 0.2	3.3 ± 0.3
G	*t*-butylglycine	0.9 ± 0.1	2.0 ± 0.2
H	Phenylglycine	2.1 ± 0.6	3.1 ± 0.1
I	Cyclohexylglycine	8.9 ± 0.7	1.4 ± 0.2
J	Dimethylcysteine	0.3 ± 0.1	0.5 ± 0.1
K	Thioproline	2.2 ± 0.4	6.3 ± 1.0
L	Arginine	2.5 ± 0.2	-

3.2.3.2.1 Characterisation of peptide J binding to VBC

The binding of peptide **J** to VBC was characterised by different biophysical techniques, in addition to BLI, and all showed consistent results (Figure 3.10). Titration of 5 mM of peptide into 100 μM of VBC by ITC (Figure 3.10B) resulted in the determination of K_d = 5.28 x 10^{-4} ± 0.65 x 10^{-4} M (average of two experiments). This three-digit micromolar K_d was concordant with the K_d value obtained by BLI. Peptide **J** was also titrated in the AlphaLISA competition assay (described in 2.2.5.2), alongside the parental peptide (**A**), to assess its ability to displace Rbx1–Cul2 from VBC (Figure 3.10C). It was observed that both peptides were able to displace Rbx1–Cul2 from VBC with an IC_{50} = 2.2 x 10^{-4} M, a surprising result, considering the low binding affinity of the peptides determined by BLI and ITC. Due to the narrow dynamic range of the experiment, this assay could not distinguish the binding of peptide **J** from peptide **A** (K_d = 2.0 ± 0.2 mM by BLI). Nonetheless, it clearly validated the binding to the Cul2 binding site. Protein-observed NMR also suggested that the peptide was binding to VBC (concentration-dependent peak shifting observed, data in Appendix I). Additionally, chemical shift perturbation (CSP) analysis corroborated the AlphaLISA experiment by suggesting that peptide **J** was binding to the EloC pocket (Figure 3.11). All the residues that shifted in the presence of the peptide were identified based on the resonance peak assignment[138] available for VBC and the distances relative to the reference peak (VBC in the absence of peptide) were calculated. Only residues that shifted more than the average plus one standard deviation ($\overline{x} + \sigma$) were considered as significantly shifted. Mapping the significantly shifted residues on the structure suggested that the residues more affected by the presence of the peptide were near the EloC pocket. There were also some other peaks shifting

that correspond to residues in different areas of the protein. It is expected that amino acids at a certain distance from the binding site might rearrange upon binding of a ligand and, hence, promote shifts. On the other hand, it was observed that some peaks corresponding to a certain set of residues were always shifted, independently of the ligand titrated (observation based on data from other experiments) and those were not taken into account.

It is anticipated that the structure of Cul2 N-terminal loop bound to VBC can differ significantly when in the context of an 8-mer peptide compared to when in the context of the full-length protein. The structural knowledge of the binding mode of the peptide in complex with VBC would help to inform molecular design to enhance binding. Considering this, efforts were taken into co-crystallising peptide **J** with VBC. Soaking of the peptide into VBC crystals was considered, however, according to previous VBC crystal structures, the peptide seemed too long to reach the binding site through the channels formed by crystal contacts from packing. Therefore, a co-crystallisation approach was taken: peptide **J** and VBC were mixed in 20:1 ratio and the mixture was used to set up sparse-matrix screenings (JCSG+ and Classics from Qiagen and ProPlex from Molecular dimensions). There was one hit condition in the JCSG+ screen, which was reproduced, however, the crystals were too small (Figure 3.10D). Despite some optimisation efforts there were no great improvements and since there was no confidence that the crystals were protein crystals this condition was not followed up.

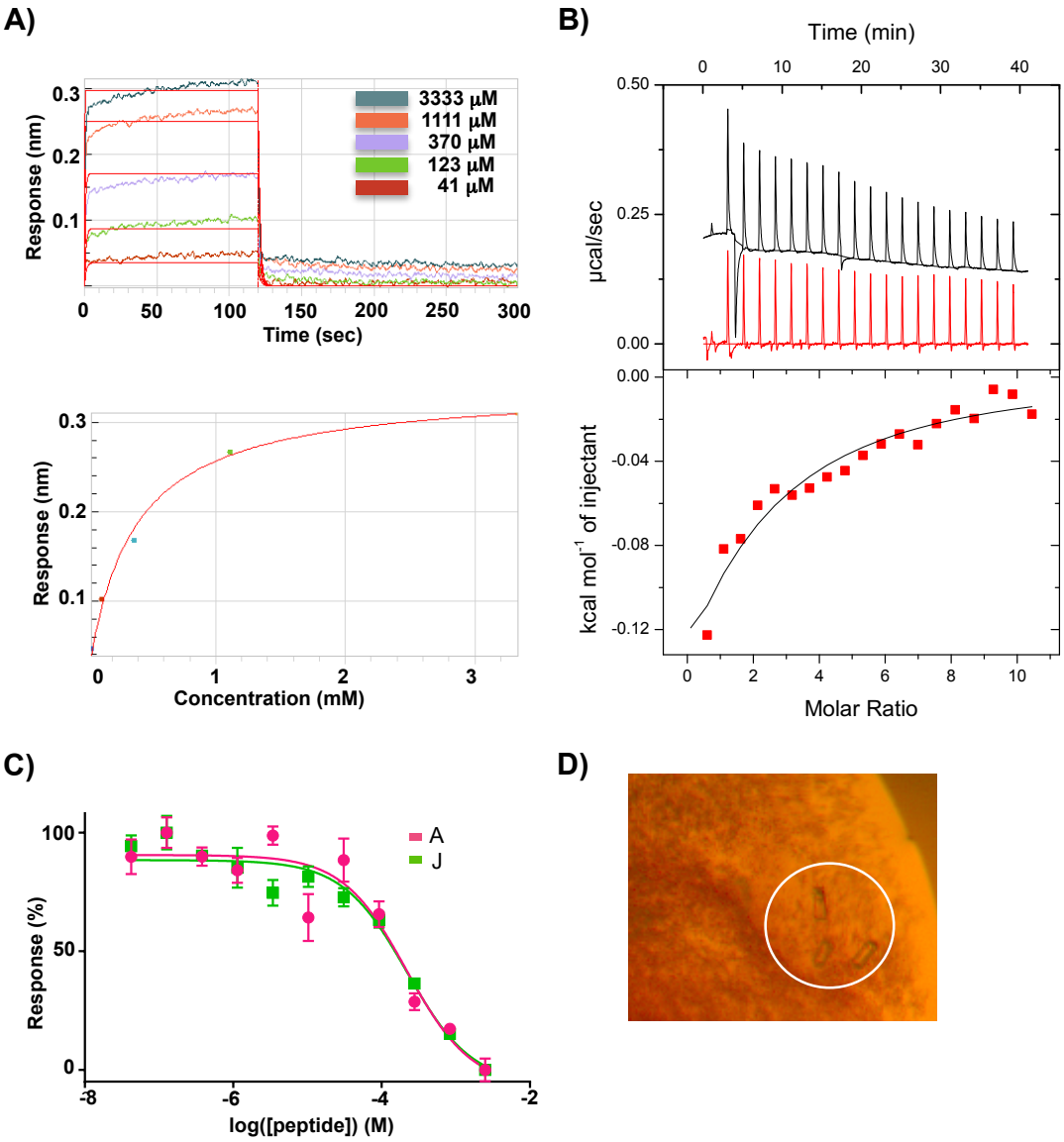


Figure 3.10 – Biophysical characterisation of peptide J binding to VBC. A) BLI results. Peptide **J** was titrated against VBC and a K_d value was determined based on the fitting of the response data points. **B) ITC experiment.** 5 mM of peptide **J** titrated against 100 µM VBC at 298 K (red); a control experiment was performed, in which peptide **J** was titrated in the same conditions against buffer (black) and this data was used to subtract the heat of dilution from the first experiment. **C)** AlphaLISA results of the titration of peptides **A** and **J** in a competition experiment to displace Rbx1–Cul2. **D)** Co-crystallisation efforts of peptide **J** with VBC. Crystal objects appeared in 0.1 M citrate pH 5.0, 3.2 M $AmSO_4$.

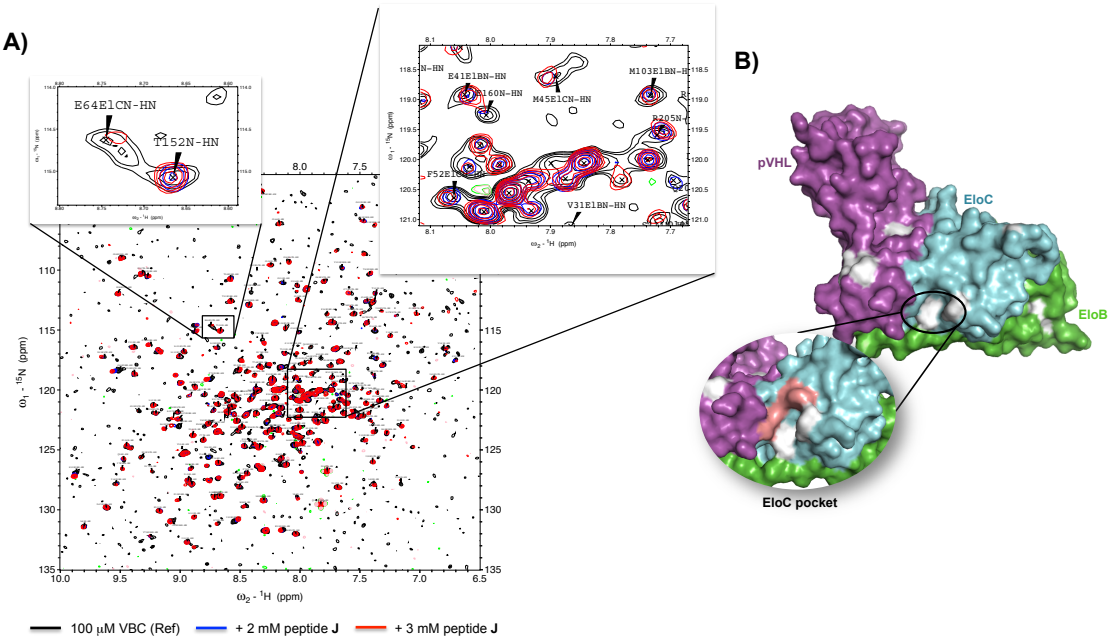


Figure 3.11 – Titration of peptide J into VBC analysed by protein-observed NMR. A) HSQC spectra with zoom in of shifted peaks. **B)** Residue mapping of the residues shifted by CSP. Residues in white are shifted more than the average plus one standard deviation ($\overline{x} + \sigma$); residues in pink are not assigned in our spectra.

3.2.4 Cyclic pentapeptides

The motivation for designing cyclic peptides arose from previous results suggesting that the conformational pre-arrangement of the peptide was required for the binding event. As seen before, the shortest peptide able to bind VBC was constituted by 6 amino acids (MSLKPR) and mutation of Pro5 or Arg6 to alanine in the context of an 8-mer peptide resulted in total loss of binding. Pro5 and Arg6 seem to play important roles in stabilising the folded conformation (Figure 3.12) and, therefore, it was hypothesised that this folded arrangement was a requirement for Cul2 peptides to bind to VBC.

It was hypothesised that a way of stabilising this structural arrangement was by cyclising the peptide, closing a ring between the arginine side chain and the serine backbone. Thus, a small library of cyclic peptides with five amino acids (Ser-Leu-Lys-Pro-X), in which the side chain of the last amino acid could be linked to the backbone of the first residue was designed, synthesised and

purified. As with the linear peptides, the leucine residue was substituted with natural and non-natural amino acids in an attempt to tighten the binding affinities (Figure 3.13). The cyclic peptides were tested for binding to VBC using BLI, DSF and ligand-observed NMR.

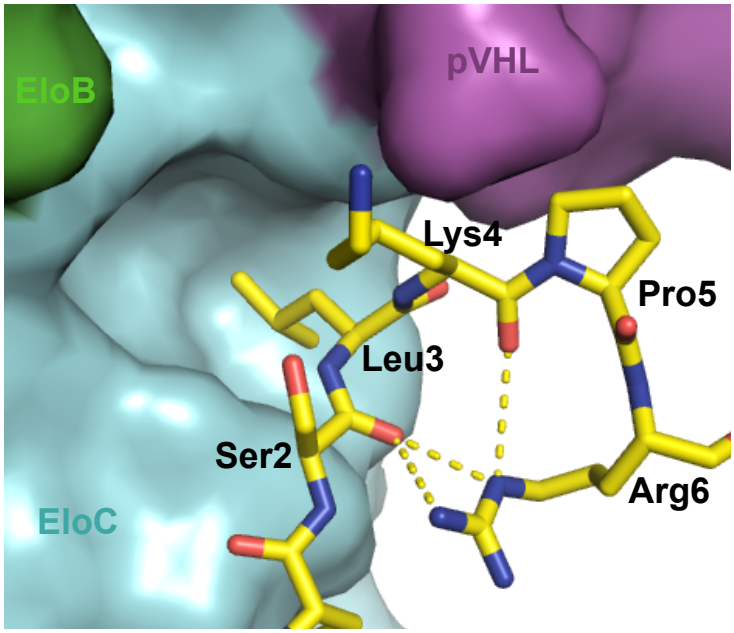
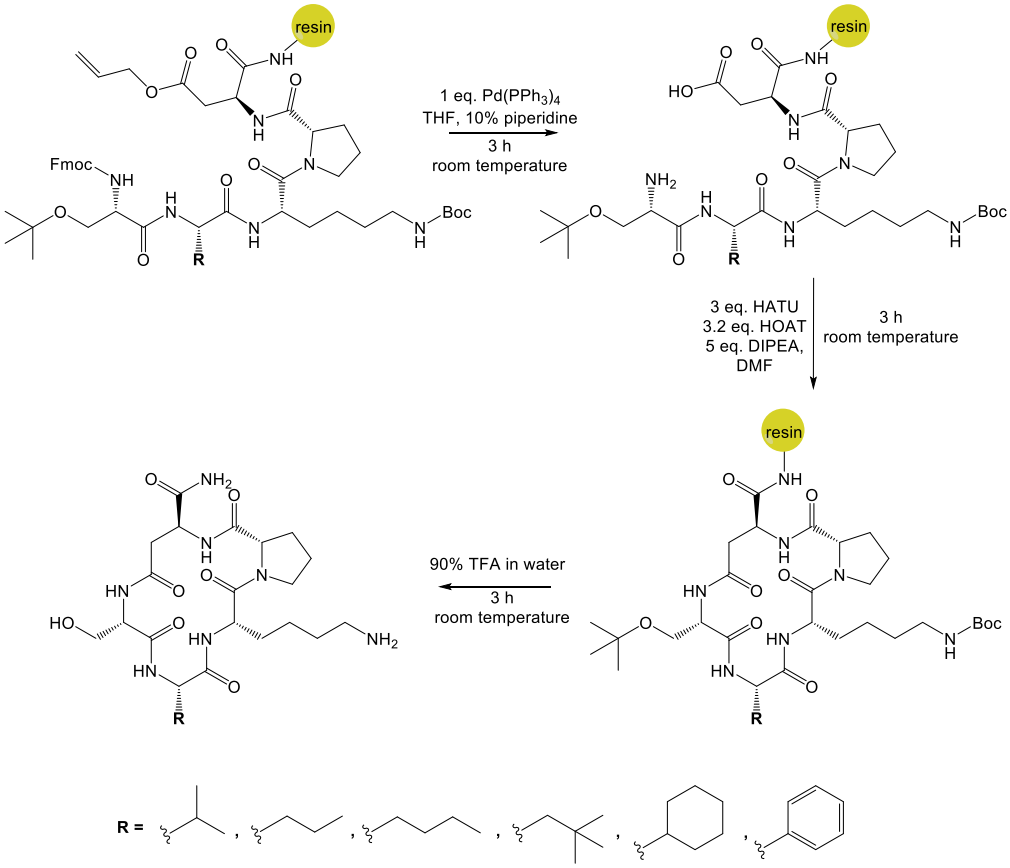


Figure 3.12 – N-terminal loop of Cul2 interacting with EloC and pVHL in the crystal structure. The first six amino acids of Cul2 present themselves in the crystal structure with a conformation that resembles a cyclic peptide. This conformation seems to be stabilised by the side chain of Arg6 that establishes hydrogen bonds with the backbone of Ser2 and Lys4.

3.2.4.1 Peptide synthesis and purification

The design, synthesis and purification of the cyclic peptides was planned and executed in collaboration with Dr. Andrea Testa. In these pentapetides, Arg6 was replaced by aspartate or glutamate bearing an O-allyl protecting group in the side chain. The linear peptides were synthesised as in previous sections, however, using a low-loading capacity resin (0.2-0.4 mmole/g), to allow the cyclisation to be performed on-resin and avoid intermolecular cyclisation. After synthesis of the linear intermediates, simultaneous deprotection of the carboxylate group of the side chains of aspartate or

glutamate and the N-terminal Fmoc was performed using stoichiometric amount of $Pd(PPh_3)_4$ and an excess of piperidine in THF. This was followed by cyclisation reaction, performed with HATU, HOAT and DIPEA, which linked the side chain of the C-terminal residue to the backbone of the N-terminal residue.[139] By using glutamate or aspartate as C-terminal amino acids, the size of the macrocycle varied by one carbon atom. Then, the cyclic peptides were cleaved from the resin and the remaining side chain protecting groups were removed using 95% TFA in water. The peptides were purified by HPLC in basic conditions and characterised by analytical HPLC-MS. These experiments were performed in parallel with the linear peptide derivatives, thus prior to the discovery of the stronger binding shown by the dimethylcysteine derivative in the position corresponding to Leu3 (corresponding to peptide **J**). For this reason, dimethylcysteine was not included in this initial library of cyclic peptides



Scheme 3.1 – Synthesis of the cyclic pentapeptides.

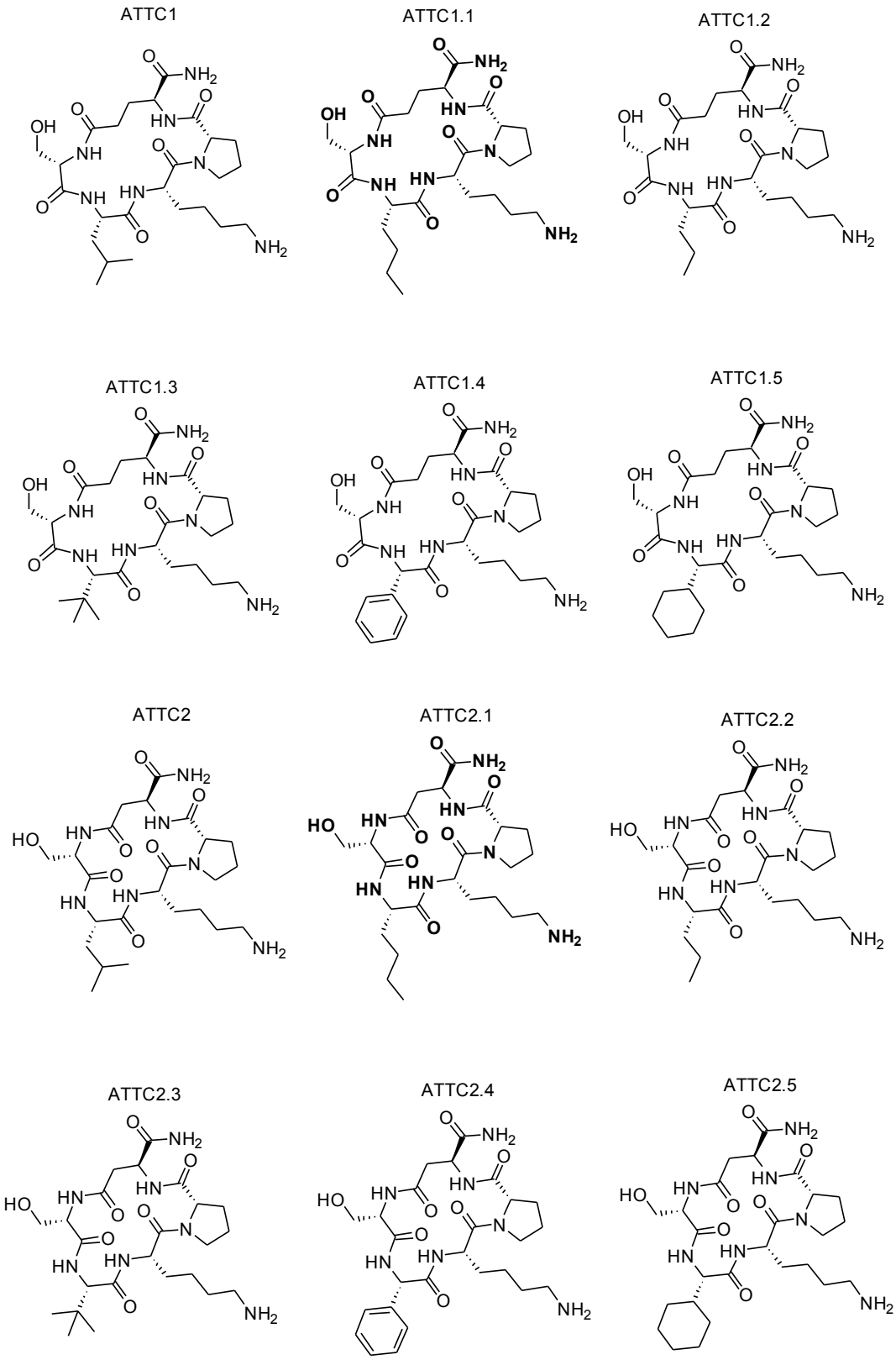


Figure 3.13 – Library of cyclic peptides. Chemical structures of the cyclic peptides containing five amino acids, different ring sizes and different side chains in place of Leu3.

3.2.4.2 Biophysical characterisation

The peptides were tested for binding towards VBC by BLI, DSF and ligand-observed NMR. The results, in some cases, showed very weak binding but in most of the cases no binding was detected (results not shown). Some of the peptides were also soaked into VBC crystals, datasets were collected from these crystals but unfortunately no peptides were found to be bound in any of these structures.

3.2.5 Peptides targeting the charge complementarity interface

The interaction in the EloC pocket is essentially driven by hydrophobic contacts and this poses a potential problem in identifying ligands due to lack of specificity. Having this in consideration, targeting another point of interaction between Cul2 and VBC was contemplated. Despite the flat characteristics of the PPI involving helix α5 of Cul2 and VBC, it seemed possible to target this interaction with peptides that could recapitulate the electrostatic charge complementarity feature (Figure 3.14A). Hence, peptides containing 10 or 15 residues of the helix α5 of Cul2 were synthesised. The synthesis and purification of these peptides were carried out as described earlier.

3.2.5.1 Biophysical characterisation

The 10-mer peptide was extremely insoluble, therefore, only the 15-mer peptide (NH_2-DCLYRYLNTQFIKKN-$CONH_2$) was tested for binding towards VBC. The peptide was first tested by DSF and it was found to promote a decrease in the T_M of VBC of about 3 °C (Figure 3.14B). Next, a ligand-observed NMR experiment (CPMG) was run and a modest decrease in the signal of the peptide was observed in the spectrum in the presence of VBC

(Figure 3.14C). This result suggested that the 15-mer peptide could bind to VBC, although it appeared to be a very weak interaction. In addition, protein-observed NMR was employed to investigate the binding site of the 15-mer peptide. The peptide was titrated at increasing concentrations into a sample of $^{2}H,^{15}N$-VBC and HSQC spectra were collected for each concentration tested. The superposition of the spectra highlighted a number of peaks corresponding to residues in VBC that were perturbed and shifting in a concentration-dependent manner in the presence of the peptide, including Lys159, a residue directly involved in the interaction (Figure 3.14D). This result hinted that the peptide was binding at the expected binding site. However, unfortunately, most of the residues at the peptide's binding site were not distinguishable in the HSQC spectra, and, therefore, the interaction site could not be confirmed. Finally, a $K_d = 1.8 \times 10^{-3}$ M was determined by BLI for the interaction of the peptide with VBC (Figure 3.14E). Altogether, the set of biophysical experiments performed suggested binding of the 15-mer peptide to VBC. However, the data also pointed to a relatively weak binding affinity.

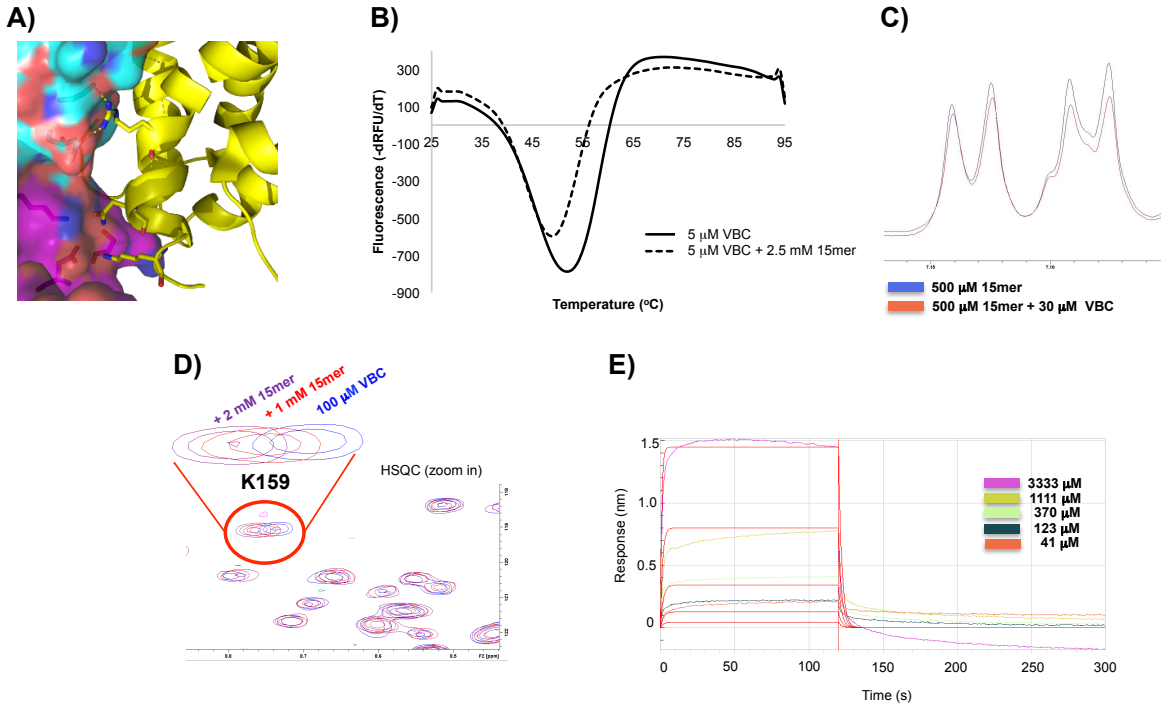


Figure 3.14 – Biophysical characterisation of the 15-mer peptide targeting the charge complementarity interface. A) Charge complementarity interface in the structure of Cul2 (yellow) bound to VBC (pVHL in purple and EloC in blue). **B)** DSF data showed a negative shift for the melting temperature of VBC in the presence of an excess of 15-mer peptide. **C)** Ligand-observed NMR data. CPMG experiment showed a reduction of the 15-mer peptide signal in the presence of VBC (red). **D)** Protein-observed NMR. Zoom in of the HSQC spectra showing the concentration-dependent shift of residue Lys159 directly involved in the interaction. **E)** BLI data of the titration of the 15-mer peptide against VBC, yielding a K_d of 1.8 x 10^{-3} M.

3.2.6 Clamp peptides

In an attempt to increase the binding affinity of the Cul2 peptides for VBC, an idea to design a peptide linking two epitopes related to the Cul2–VBC interaction emerged (Figure 3.15). The two epitopes are constituted by a 6-mer peptide of the Cul2 N-terminus and an 8-mer peptide that should recapitulate the electrostatic interactions between Cul2 and pVHL. The goal was to increase the tightness of the interaction by increasing the points of contact between the peptide and VBC. Thus, a flexible linker composed of glycine, serine and lysine residues was designed considering the distance between the two attachment points.[140] This distance, in the crystal structure, is of 18.8 Å. From analysis of the $Cul2_{1-163}$-VBC crystal structure,[82] the average distance covered by an amino

acid in a peptide without secondary structure (amino group to amino group of the next residue) is of 3.2 Å, thus the linker was designed to have a fully extended length between the two epitopes larger than the distance measured in the crystal structure. Two peptides were designed, one containing the original leucine and a second one in which the leucine was replaced by dimethylcysteine, as in peptide **J**. These peptides, designated as clamp peptides, were synthesised and purified and their binding towards VBC was evaluated. The peptide synthesis and purification was performed according to the methods already described.

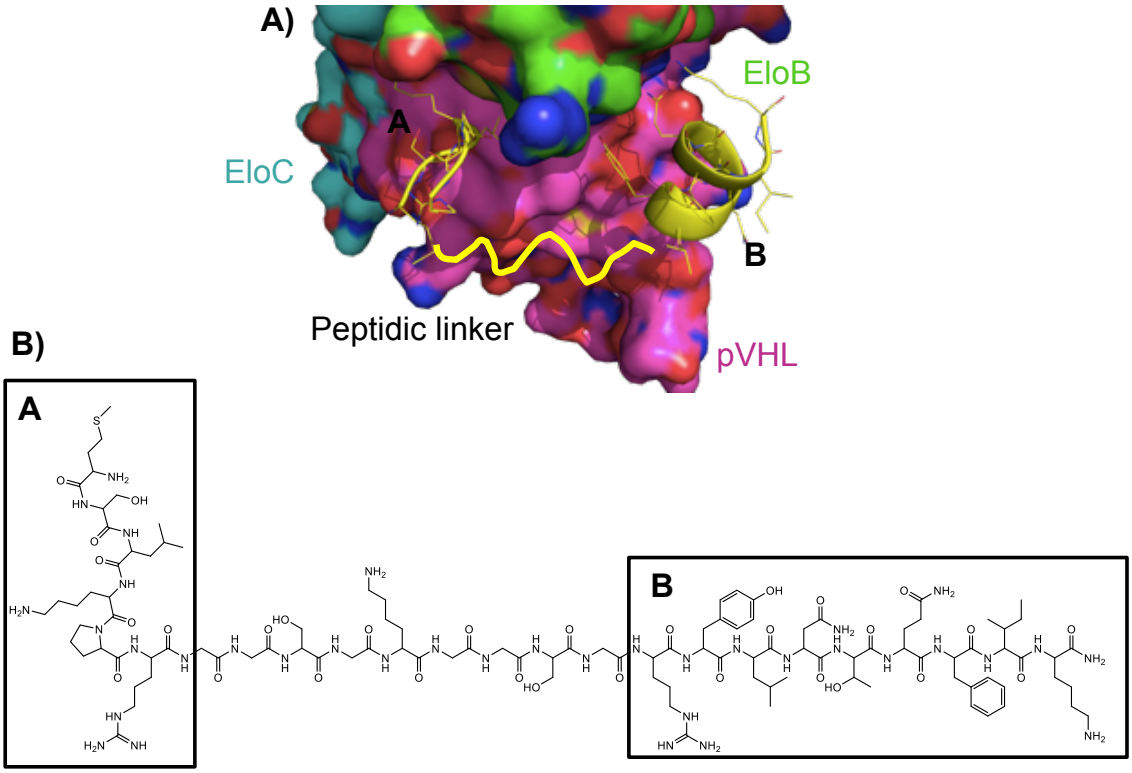


Figure 3.15 – Clamp peptide. A) Illustration of the peptidic linker designed to join the two epitopes in the structure in the crystal structure of Cul2–VBC. **B)** Chemical structure of the clamp peptide A.

3.2.6.1 Biophysical characterisation

The clamp peptides were very insoluble at high concentrations, which constituted an obstacle for the biophysical characterisation. The only reliable results were obtained in the AlphaLISA competition assay (Figure 3.16) where no significant difference was observed between the two peptides, clamp **A** and clamp **J**. However, in this assay it was observed a 6-fold decrease in the IC_{50} in comparison to the parental peptide **J** (36 μM vs. 220 μM, respectively) suggesting a greater ability of the clamp peptide **J** to displace Cul2 from VBC.

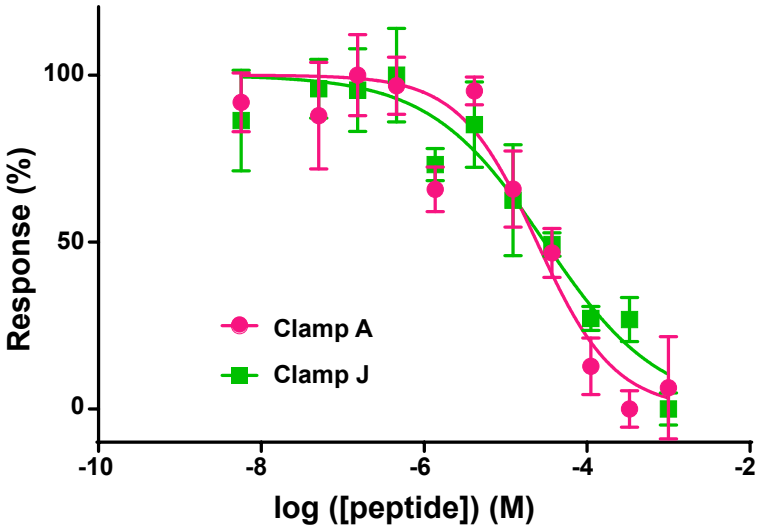


Figure 3.16 – Dose-response curves of the AlphaLISA data showing the ability of the clamp peptides to displace Cul2. Clamp peptides were titrated into a mixture of bead-bound VBC and Rbx1–Cul2. The experiments were performed in quadruplicate and the data points shown are the average values for each concentration. The error bars represent the standard deviation of each point.

Efforts were conducted to obtain a crystal structure of either of the clamp peptides bound to VBC and use it to guide ligand optimisation. Co-crystallisation experiments were set up, however, no hits were identified. The low solubility of these peptides associated with the low binding affinity to VBC hindered the identification of suitable crystallisation conditions, and they were not pursued further.

3.3 Discussion

Probing the interaction surface between VBC and Cul2 with peptides carries great potential and opportunities for the development of modulators of this interaction. Previous work in the Ciulli laboratory has led to the development of potent small-molecule disruptors of the pVHL–HIF-1α PPI based on the structure of pVHL bound to a 19-mer parental peptide derived from HIF-1α.[43,52] It was therefore deemed as an attractive strategy to explore the potential to apply a similar approach to other non-HIF binding surfaces on VBC. In addition, the identification of fragments binding to a new pocket on the EloC subunit suggested that this surface on VBC could be ligandable, thus exploring avenues to target it with peptides was of interest. This chapter described efforts conducted towards the ultimate goal of developing a chemical probe for this PPI.

A construct that allows easy expression of Cul2 with a high yield and in an active and stable form would be of great advantage for the biophysical evaluation of potential inhibitors of the Cul2–VBC interaction and would allow the development of robust assays to screen modulators of this PPI. The merging of these two parallel tactics could aid in the investigation of Cul2–VBC interaction and would help to determine whether targeting this site provides effective approaches and chemical probes to interrogate CRLs biology and modulate their activity.

Protein aggregation is a major issue in recombinantly-expressed proteins. Aggregation can be caused by the intrinsic properties of certain proteins but it can also be related to factors affecting the expression and purification conditions. The co-expression of MBP-tagged $Cul2_{NTD}$ with chaperones yielded significant amounts of soluble $Cul2_{NTD}$, however, it was

shown that the protein was extremely prone to aggregation, existing as a mixture of multiple species. The one-off experiment where monomeric species of $Cul2_{1\text{-}384}$-MBP-6xHis was obtained in a considerable amount hinted the aggregation could be promoted by the binding to the MBP trap. However, this hypothesis was rapidly discarded as it was observed that the monomeric species was unstable, forming aggregate as soon as it was isolated. Despite the different constructs, tags, expression conditions, lysis methods, purification methods and solubilisation efforts, it was not possible to obtain active, stable and monomeric $Cul2_{NTD}$ using *E. coli* as expression system. In fact, there is no report of soluble Cul2 expressed in *E. coli*, with exception of the work from Nguyen *et al*.,[82] where $Cul2_{1\text{-}163}$ is co-expressed with VBC. Perhaps this was an ambitious goal with high risk-benefit balance, that albeit frustrating provided evidence that certain human proteins simply require an eukaryotic expression system to yield functional recombinant proteins.[141,142]

Cul2-derived peptides binding to the EloC site could provide valuable insights on how to target the Cul2–VBC PPI. The biophysical results indicated that at least six residues (MSLKPR) were needed to observe a binding event and the experiments performed allowed to confidently estimate a K_d value of approximately 2 mM for the interaction between an N-terminal 8-mer peptide of Cul2 and VBC. In addition, Pro5 and Arg6 residues play an important role in the binding of the peptide, as the respective mutations to alanine resulted in complete loss of binding. This observation is supported by the crystal structure where the arginine residue seems to fold back and establishes intramolecular hydrogen bond contacts that stabilise the peptide fold, which seems to be induced by the proline residue.

Replacing Leu3 with a set of natural and non-natural amino acids confirmed that side chains bulkier than leucine's are tolerated for binding. There was no significant improvement in the binding affinity for any of the peptides tested with the exception of peptides **G** and **J.** In these two cases, the leucine residue was replaced by a *t*-butylglycine and a dimethylcysteine, respectively, and the K_d values determined by BLI dropped below the millimolar range (~ 900 μM for peptide **G** and ~ 300 μM for peptide **J**). The biophysical characterisation of peptide **J** suggested moderate binding to VBC and it disclosed the opportunity to develop Cul2 peptides into high-affinity binders. The strength of the interaction was boosted about 4-fold simply by increasing the volume of the side-chain fitting the EloC pocket. Determining the binding pose of peptide **J** would facilitate the structure-based design and help the optimisation of the ligand. If this is not achievable by X-ray crystallography in the future, perhaps, computational or NMR methods could be applied to predict the binding mode and guide ligand design.

Structural pre-arrangement of peptides could enhance binding affinities by reducing the entropic penalty upon binding. A library of cyclic pentapeptides was conceived based on this consideration and based on the conformation of the N-terminal tail of Cul2 observed in the Cul2–VBC crystal structure. Unfortunately, weak or no binding affinities were observed, which could potentially be explained by suboptimal cyclisation constraining the peptides or it could also be related to the purely hydrophobic nature of the interaction. Considering the lack of specificity conferred by a merely hydrophobic interaction, other PPI contacts were explored. The 15-mer peptide targeting the helix α5-VBC interaction proved to be a very weak binder. Perhaps a way of tightening the interaction could be to stabilise the helical conformation of the

peptide by peptide stapling,[61] for example, mimicking the original PPI. On the other hand, the clamp peptides linked two epitopes of Cul2 and promoted a 6-fold reduction in the IC_{50} value compared to the parental 8-mer peptide **J**. There were issues with the solubility of these peptides that could be addressed in the future by replacing the linker with a more hydrophilic chain, for example. Once these peptides are soluble at higher concentrations, a proper biophysical assessment can be performed. From then, structural information on the binding mode could help ligand-optimisation.

In conclusion, it has proven very challenging to target the PPI between VBC and Cul2, potentially because of its hydrophobic driven and relatively flat features (the buried surface area is 1022 $Å^2$), which might compromise the development of high affinity ligands. Other drug discovery tools such as peptide stapling,[61] tethering,[55] or even helix mimetics,[62] for example, might be helpful toward this goal. Another approach to target this kind of PPIs could be the use of bicyclic peptides as they cover a large surface area and are able to closely mimic PPI features.[143] In a later stage, combining the information obtained from the fragment work with the peptide work could lead to the design of merged molecules that could act as potent chemical probes to study CRLs biology.

CHAPTER 4

BICYCLIC PEPTIDES AND STUDIES OF THEIR BINDING TOWARDS THE VBC COMPLEX

4.1 Introduction

Cyclic peptides are considered an attractive modality for the development of chemical tools and molecular therapeutics as they can combine key features such as good binding affinity, target selectivity and low toxicity.[63] However, the application of cyclic peptides to probe protein surfaces is relatively under explored. The term bicyclic peptide used in this chapter refers to cyclic peptides containing two relatively flexible loops attached onto a rigid central scaffold. Bicyclic peptides tend to have a more constrained structure, when compared to monocyclic or linear ones. This arrangement allows a closer mimicry of antibodies in terms of their molecular recognition, binding affinity and specificity properties (Figure 4.1).[144]

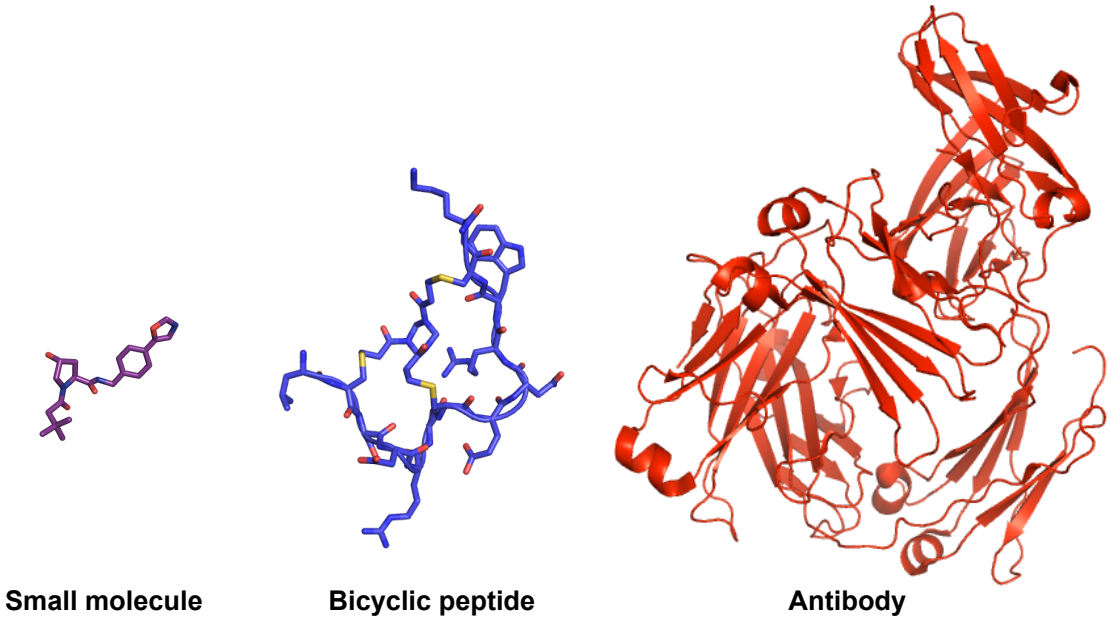


Figure 4.1 – Differences between a small-molecule, a bicyclic peptide and an antibody. The small-molecule (purple) is a VHL inhibitor (PDB: 4W9C)[43]; the bicyclic peptide (blue) is a inhibitor of the urokinase-type plasminogen activator (PDB: 4MNX)[145]; and the antibody (red) is the human IgE-Fc (PDB: 5LGK).[146]

In addition, these structural features have been shown to introduce improvements in binding affinities as well as resistance to intracellular degradation and metabolic activities.[147] Hence, bicyclic peptides are suitable

candidate probes to modulate protein-protein interactions as they are able to establish extensive contacts (buried surface areas > 700 Å)[148] with globular proteins and they are expected to bind tightly to protein surfaces.

Heinis and Winter[149] have described a strategy to select bicyclic peptide ligands, which combines the phage display technology with the covalent attachment of a chemical core onto a linear peptide. Phage display is an *in vitro* screening technique for the identification of ligands for proteins and it entails the expression of a peptide sequence as fusion to the coat protein of a bacteriophage. Thereby the phage-displayed peptides are physically linked to their encoding DNA, permitting the selection of binding partners by iterative rounds of biopanning and amplification, followed by DNA sequencing.[150] This is one of the major advantages of this technology, together with the possibility of screening over a billion different molecules at the same time. The ultimate goal of a phage display screening is to identify a consensus sequence or motif that is enriched throughout the rounds of biopanning, consistent with a tight, target-specific binder.

In this chapter, the application of phage-display technology to the screening of bicyclic peptide libraries against VBC is described. The libraries used were obtained from Prof. Christian Heinis (EPFL, Lausanne). The peptides displayed at the N-terminus of the minor coat protein pIII of the phage contained three reactive cysteine residues (spaced apart by a variable number of random amino acids), which allowed conjugation with 1,3,5-tris-(bromomethyl)benzene (TBMB).[143] As a result, a peptide formed by two peptidic loops linked by a mesitylene core is generated (Figure 4.2) and screened against the target of interest. Each phage particle presents five copies of the pIII protein, therefore, five peptides are displayed on its surface. The libraries were constructed using

degenerate primers to code for the randomised regions of amino acids, in between the conserved cysteine residues.[151]

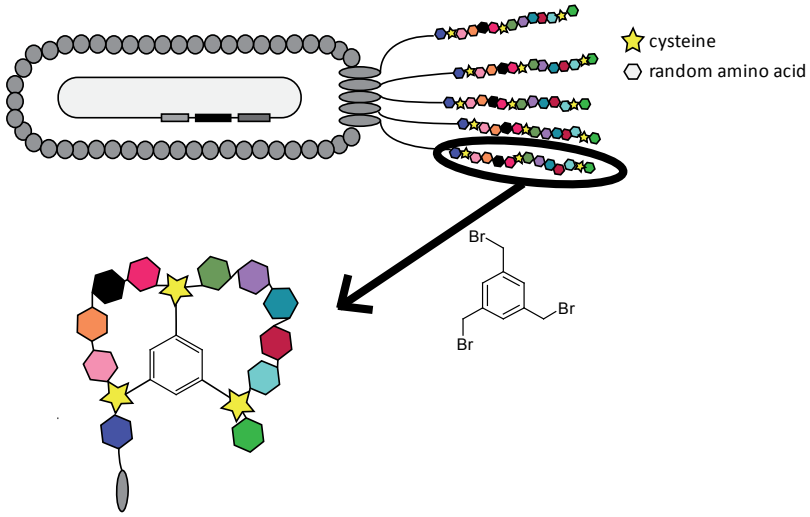


Figure 4.2 – Phage display of bicyclic peptides. Representation of the peptides displayed on the phage coat being cyclised with TBMB prior to screening.

The rationale was, therefore, to identify peptide binders of the VBC complex using an unbiased approach. It was intended to understand whether bicyclic peptides could be accommodated in either the HIF-1α site or the new binding sites identified by fragment screening from previous work in the Ciulli laboratory, and whether other binding sites within the complex could be identified. In this chapter, the results of a phage display screening of three libraries of bicyclic peptides are presented and the respective results are discussed.

4.2 Results

4.2.1 Phage display screening

Three phage-display libraries of bicyclic peptides were screened against VBC and VBCH (Table 4.1). The differences between each library lay on the number of randomised amino acids that constitute each loop. The reason to perform the screening against VBC and VBCH was to potentially discriminate between ligands of the VBC–HIF-1α interface and of other binding sites as a mean to aid their identification. The size of each library was determined experimentally by counting the number of bacterial colonies formed after transformation.[143]

Table 4.1 – Phage display libraries of bicyclic peptides. VL1, VL2 and VL3 libraries were screened. The library size, expected bicyclic format according to possible loop lengths [$(X)_m x (X)_n$] are presented. The library size was experimentally determined.[143]

Library	Bicyclic peptide format	Loops	Library size
VL1	$AC(X)_mC(X)_nCG$ m + n = 7 or 8	3x4; 4x3; 4x4; 3x5; 5x3	5×10^8
VL2	$AC(X)_mC(X)_nCG$ m + n = 9	3x6; 6x3; 4x5; 5x4	1×10^7
VL3	$AC(X)_mC(X)_nCG$ m + n = 10 or 11	4x6; 6x4; 5x5; 5x6; 6x5	6×10^7

Briefly, the screening protocol[152] (Figure 4.3) consisted first in the amplification of the library using *E. coli* as host organism. The amplification is an important step to ensure that each member of the library is represented at least 10 times, on average.[151] Then, the amplified phages were isolated from the bacteria and the peptides displayed on the phage coat were cyclised by incubating a diluted phage solution with excess of TBMB under reducing

conditions (1h at 42 °C). Next, in a step named biopanning, the phages were screened against the target, which had been prior immobilised covalently on tosyl-activated magnetic beads (Dynabeads M-280 tosyl-activated, Invitrogen). The phage solution of bicyclic peptides was incubated with the magnetic beads loaded with VBC or VBCH for 1h at room temperature, allowing peptides to bind to the protein. After this step, the beads were washed thoroughly to dissociate the unbound phages and then the phages that were still attached to the protein (binders) were eluted with a solution of glycine at low pH. In the next step, the process was repeated from the beginning, starting with the amplification of the phages from the previous round. In total, four rounds of biopanning were performed for the screening of each library, in an attempt to enrich the binders. The DNA of thirty clones from the last round of biopanning of each library was sequenced and the peptide amino acid sequences were analysed.

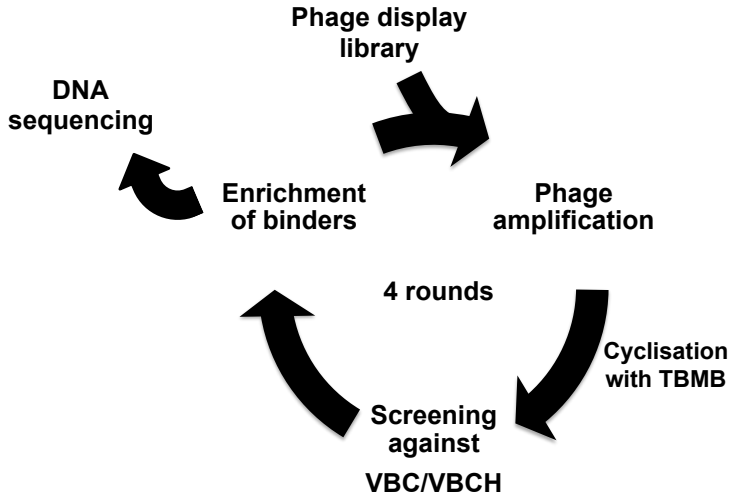


Figure 4.3 – Phage display screening methodology applied. The methodology applied consisted on 4 rounds of biopanning, which contemplated phage amplification, cyclisation with TBMB, screening against the immobilised target and enrichment of the binders.

Throughout the rounds of biopanning, a negative control was performed in parallel, where the phage solution was incubated with the beads in the

absence of protein. It was expected that the number of colonies obtained in the negative control would be lower than the number of colonies from the biopanning round, which was indeed observed in all cases.

4.2.1.1 Screening of the VL1 library

The VL1 library had a complexity of 5 x 10^8, the number of random amino acids varied between 7 and 8, and the loops forming the bicyclic peptides could assume 5 different formats, (3x4, 4x3, 4x4, 3x5 or 5x3). The screening of this library in the format of bicyclic peptides led to the identification of 19 different peptide sequences against VBC and another 19 against VBCH (Table 4.2). Surprisingly, there were no common sequences between the VBC and VBCH screening. Some sequences of the peptides identified were observed more than once but the sequences within each of the screenings (VBC or VBCH) were very variable and no consensus motif could be identified.

Table 4.2 – Screening of the VL1 library. Sequences of peptides pulled down after the fourth round of biopanning against VBC (left panel) and VBCH (right panel). The number on the left of each peptide reflects the number of occurrences of the sequence.

	VL1 vs. VBC	
1	AC**C**QEAC**C**DIWPC**C**G	
1	AC**C**IIWC**C**TLIFC**C**G	
2	AC**C**WSYC**C**PTYIC**C**G	3x4
1	AC**C**QMVC**C**QHTIC**C**G	
2	AC**C**SYYC**C**SWTQC**C**G	
1	AC**C**GNTC**C**ANLSIC**C**G	
1	AC**C**APWC**C**YILWHC**C**G	
1	AC**C**PQTC**C**LVSHRC**C**G	3x5
5	AC**C**AYSC**C**VLQGIC**C**G	
1	AC**C**QDLC**C**WWSNLC**C**G	
4	AC**C**SISMC**C**NVTC**C**G	4x3
1	AC**C**MRSIC**C**TDGTC**C**G	
2	AC**C**SNYHC**C**PQHEC**C**G	4x4
1	AC**C**IRHTC**C**ELSNC**C**G	
2	AC**C**LDLRC**C**QLPYC**C**G	
1	AC**C**LFVLTC**C**LFYC**C**G	
2	AC**C**PFGVDC**C**QYLC**C**G	5x3
1	AC**C**SSVQFC**C**GHLC**C**G	

	VL1 vs. VBCH	
2	AC**C**PRMC**C**HTLPC**C**G	
1	AC**C**VPQC**C**GDPQC**C**G	
1	AC**C**TTEC**C**AVKQC**C**G	3x4
5	AC**C**RARC**C**PLEQC**C**G	
1	AC**C**VNNC**C**HAPIC**C**G	
1	AC**C**AVQC**C**HFPHC**C**G	
1	AC**C**VPGC**C**ATMCAC**C**G	3x5
2	AC**C**VLFC**C**FFLLCC**C**G	
1	AC**C**AGANC**C**INSC**C**G	
1	AC**C**QTTVC**C**HFDC**C**G	4x3
3	AC**C**LPRQC**C**QDLC**C**G	
1	AC**C**TNLTC**C**SSHC**C**G	
1	AC**C**PALSC**C**SASNC**C**G	
2	AC**C**QHRKC**C**YAFVC**C**G	
1	AC**C**WDVRC**C**SQLFC**C**G	4x4
1	AC**C**LQHAC**C**MDIRC**C**G	
2	AC**C**ASRTC**C**NVLSC**C**G	
2	AC**C**HSTKSC**C**RNSC**C**G	5x3
1	AC**C**TQRNHC**C**APLC**C**G	

4.2.1.2 Screening of the VL2 library

The VL2 library comprised 14-mer peptides, with nine of the fourteen residues random, and had a complexity of 1 x 10^7. The loop sizes and arrangements of these bicyclic peptides could vary between 3x6, 6x3, 4x5 or 5x4. The biopanning of this library cyclised with TBMB generated 21 different peptide sequences against VBC and 20 against VBCH out of 30 sequencing results in each case (Table 4.3). This time, there were five peptide sequences in common between the screening of VBC and VBCH.

Table 4.3 – Screening of the VL2 library. Sequences of peptides pulled down after the fourth round of biopanning against VBC (left panel) and VBCH (right panel). The number on the left of each peptide reflects the number of occurrences of the sequence. The sequences in red were simultaneously pulled down against VBC and VBCH.

	VL2 vs. VBC	
1	ACGPVCPTSQLLCG	
4	ACEERCPMAQSGCG	3x6
1	ACLRGCQNAQTTCG	
1	ACGLRGCQHTEFCG	
1	ACIQSVCYQGSKCG	
1	ACFDQLCWFYGICG	
1	ACDLRACQHNEPCG	
3	ACQRTMCLLPQFCG	
1	ACQLQPCHGIPKCG	4x5
1	ACWHNSCYSSGQCG	
1	ACWKVSCPSSRNCG	
1	ACPSQSCEFLRMCG	
1	ACPSAACSQGIVCG	
1	ACPGAFCPKLNTCG	
1	ACSNSGTCRGSYCG	
1	ACQNSGTCRSAYCG	
1	ACSQPSMCRIMACG	
2	ACYPPRLCRSASCG	5x4
4	ACDPASRCRQDPCG	
1	ACGFTLGCAQQQCG	
1	ACSNSGTCRSGYCG	

	VL2 vs. VBCH	
1	ACQWTCSWFTCNCG	
1	ACQWTCSRFTCSCG	
3	ACEERCPMAQSGCG	
2	ACEERCPMAYSGCG	3x6
1	ACDERCPMAYSGCG	
1	ACPRICSQQSSACG	
1	ACFHYCVFFLCYCG	
1	ACHWMPCYTPQSCG	
1	ACEDHFCTVYTPCG	
1	ACPLSPCPDIRSCG	
1	ACPVKTCPPAMNCG	4x5
1	ACRFSTCNPQLECG	
5	ACFDQLCWFYGICG	
1	ACSSVPCQLAALCG	
2	ACGLRGCQHTEFCG	
1	ACGLRRGCQHTEFCG	5x5
1	ACQSSVVCSGSHCG	
2	ACSNSGTCRSGYCG	5x4
2	ACDPASRCRQDPCG	
1	ACLLPHTSCQVECG	6x3

Additionally, it was also observed some similarity between sequences within the same target. For instance, the 'NSGTCR' motif was found in three sequences in the bicyclic peptide format 5x4 against VBC; the motif 'QWTCS' was found in

two sequences of the 3x6 format against VBCH and the motif ‘EERCPMA’ was found in three peptides of the VBCH screening. Despite the still large variability of the peptide sequences pulled down, these results were encouraging.

4.2.1.3 Screening of the VL3 library

The VL3 library contained 15-16-mer peptides with 10-11 random amino acids. This third library had a complexity of 6 x 10^7 and the loops of the peptides could acquire one of following configurations: 4x6, 6x4, 5x5, 5x6 or 6x5. The screening of the cyclised library against VBC resulted in a remarkable enrichment of the sequence ACQLKRLNCHNSPLCG that was found in 20 peptides out of the 30 sequenced. The remaining two sequences identified, ACGHYKSCAYNQFCG and ACSNTRQCPSYHLRCG, were found 7 and 3 times, respectively. However, none of these sequences was observed amongst the 24 peptides found in the screening against VBCH (Table 4.4). The sequence with the highest level of enrichment for VBCH was ACLKSPLCAQAPQACG that was found 4 times and differs only in two amino acid residues from the fourth sequence of the 5x6 format.

Table 4.4 – Screening of the VL3 library. Sequences of peptides pulled down after the fourth round of biopanning against VBC (left panel) and VBCH (right panel). The number on the left of each peptide reflects the number of occurrences of the sequence.

	VL3 vs. VBC	
7	A**C**GHYKS**C**AYNQF**C**G	5x5
3	A**C**SNTRQ**C**PSYHLR**C**G	5x6
20	A**C**QLKRLN**C**HNSPL**C**G	6x5

	VL3 vs. VBCH	
1	A**C**TLLS**C**QSLHRT**C**G	4x6
1	A**C**MSQS**C**GTYSAY**C**G	
2	A**C**AQVS**C**SQRQQL**C**G	
1	A**C**PKSL**C**NQSMSL**C**G	
1	A**C**LRQT**C**VQSESI**C**G	
1	A**C**WKGS**C**RTSQSF**C**G	
2	A**C**FFYGL**C**LVFIGH**C**G	5x6
1	A**C**SQHWA**C**PMSSGS**C**G	
1	A**C**NRYVR**C**ASNSAF**C**G	
1	A**C**LKSPL**C**AYAPQG**C**G	
4	A**C**LKSPL**C**AQAPQA**C**G	
1	A**C**LSVIS**C**IGTQP**C**G	5x5
1	A**C**GKARF**C**FGNAL**C**G	
1	A**C**QLSSQ**C**TNQNN**C**G	
1	A**C**LAALK**C**SQSSA**C**G	
2	A**C**ALQVG**C**ELAPM**C**G	
1	A**C**QAVPP**C**MLTQI**C**G	
1	A**C**PKQGM**C**KLNAS**C**G	
1	A**C**MPQAL**C**SLYGT**C**G	
1	A**C**LHRAP**C**LLMQH**C**G	
1	A**C**LKAPL**C**AQAPA**C**G	
1	A**C**SHRDVN**C**PIEL**C**G	6x4
1	A**C**EMFRPP**C**THRDI**C**G	6x5
1	A**C**RQLVWS**C**GERML**C**G	

4.2.1.4 Screening of VL1, VL2 and VL3 together

The results obtained from the independent screening of each library did not yield a convergent consensus sequence and were, in fact, quite variable. Therefore, it was decided to pull together the phages from the last round of each screening (separately for VBC and VBCH) and perform an ultimate round of biopanning. The idea was that the best binders of each library would be competing with each other and the best binders amongst the three libraries would, in principle, be enriched. The results showed 7 different peptide sequences being enriched against VBC and 9 against VBCH (Table 4.5) out of the 10 clones sequenced for each screening. There were two sequences in common between the biopanning against VBC and VBCH but no general

consensus was observed, however, peptides with the format 4x5 or 5x4 seemed to be preferred in regards to others in the screening against VBC.

Table 4.5 – Screening of VL1, VL2 and VL3. Sequences of peptides pulled down after the round of biopanning against VBC (left panel) and VBCH (right panel). The number on the left of each peptide reflects the number of occurrences of the sequence. The sequences in red were simultaneously pulled down against VBC and VBCH.

	VBC	
1	ACLYTCESIFLKCG	3x6
1	ACEERCPMAQSGCG	
2	ACGLRGCQHTEFCG	4x5
1	ACFDQLCWFYGICG	
1	ACYPPRLCRSASCG	5x4
3	ACDPASRCRQDPCG	
1	ACSNTRQCPSYHLRCG	5x6

	VBCH	
2	ACEERCPMAQSGCG	3x6
1	ACQSSVVCSGSHCG	5x4
1	ACDPASRCRQDPCG	
1	ACPKQGMCKLNASCG	5x5
1	ACPTQDPCLSSIECG	
1	ACQPTSPCMAPLLCG	
1	ACPKSLNCNQSMSLCG	5x6
1	ACFFYGLCLVFIGHCG	
1	ACYNAKMSCLDPYCG	6x4

In a final attempt to improve the output of the phage display screening, phages from all the steps of the screening (round 1, 2, 3 and 4 from VL1, VL2 and VL3) against VBC and VBCH were screened together against VBC in one last round. Only 10 colonies were sequenced and the results showed that instead of obtaining a convergent sequence, as expected, 7 different peptide sequences were enriched (Table 4.6).

Table 4.6 – Last round of biopanning against VBC. Sequences of peptides pulled down after the last round of biopanning against VBC. The number on the left of each peptide reflects the number of occurrences of the sequence.

	VBC	
2	ACLYTCESIFLKCG	3x6
3	ACGLRGCQHTEFCG	4x5
1	ACPTQDPCLSSIECG	6x5
1	ACQPTSPCMAPLLCG	5x5
1	ACYNAKMSCLDPYCG	6x4
1	ACYPPRLCRSASCG	5x4
1	ACFEMTCPKQSFLCG	4x6

The phage display screening did not yield a unique consensus motif but, instead, it generated a diverse set of bicyclic peptides with different combinations of ring sizes and sequences. Some of the sequences were heavily enriched though as it was observed in the VL3 screening against VBC, where 20 out of 30 colonies shared the same peptide sequence and this was encouraging. On the other hand, the screening of the VL2 library resulted in a number of hits conserved between VBC and VBCH. Despite the overall result being different from the expected, it was hoped that the different binding motifs could provide suitable starting points for optimisation of the peptides into potent binders.

4.2.2 Synthesis of the bicyclic peptides identified from the phage display screening

Considering the diversity of bicyclic peptides obtained, to increase the chance of finding real binders of the VBC complex, all the peptide sequences identified in the phage display screening were synthesised for further evaluation. In total, 126 bicyclic peptides were synthesised. The synthesis of the bicyclic peptides contemplated two steps: synthesis of the linear peptide and cyclisation with TBMB. The linear peptides were synthesised in 15 μmole scale and the work up and cleavage from the resin were performed by mixing the resin in a solution of TFA:H_2O:1,2-ethanedithiol:phenol:thioanisole (v/v, 90:2.5:2.5:2.5:2.5) for 3h at room temperature. Next, the TFA mixture containing the peptides was filtered and the linear peptides were precipitated with cold diethyl ether. This precipitate was filtered, washed with diethyl ether, dissolved in a water:acetonitrile solution and lyophilised. For the cyclisation of the peptides, the protocol followed consisted of an alkylation reaction achieved

by the dilution of the peptide in 40 mM NH_4HCO_3 pH 8.0 in the presence of 1 mM of TBMB for 1h at 30 °C, consistent with the protocol described by Rebollo *et al*.[153] At this point it was observed that most of the peptides were not completely soluble in these conditions, as the formation of a white thin solid was observed. Even in the cases where the peptides were soluble in acetonitrile, a white precipitate was observed upon addition of the buffer.

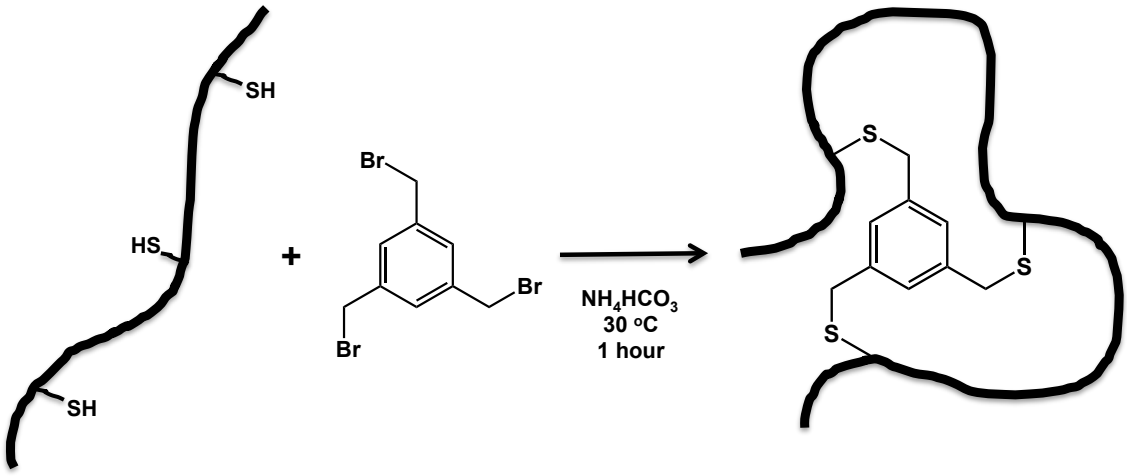


Figure 4.4 – Schematic reaction of the cyclisation of the linear peptides with TBMB to originate bicyclic peptides. The reaction is carried out in a diluted solution of peptide in 40 mM $(NH_4)HCO_3$ pH 8.0 buffer in the presence of 1 mM of TBMB for 1h in a 30 °C water bath, according to previous literature.[153]

The reaction mixture was then lyophilised and the next step was the purification of the peptides by HPLC. The purification of the bicyclic peptides was optimised on an analytical scale (Zorbax SB300 C18 4.6x250 mm column) for some of the bicyclic peptides and the optimal conditions for the isolation of the products of interest were identified. However, the efforts to scale up this purification step to the preparative range and to develop an automated method for the cyclisation-purification have not been completely successful as it was observed that the same conditions could not be applied to all the peptides due to differences in number of amino acids, total charge and polarity. Having this in

consideration it was decided to use the crude mixture for preliminary biophysical evaluation, assuming the cyclised peptide was present as the major species in the mixture. The success of the cyclisation reaction was assessed by MALDI-TOF mass spectrometry and in total, 86 peptides were confirmed to be cyclised.

4.2.3 Initial screening of the bicyclic peptides

Considering the 15 µmole scale at which the peptides were synthesised, assuming an average yield of 60% (based on a sample of 10 peptides), a molecular weight of 1500 Da and a quantitative cyclisation reaction, the crude mixture was dissolved in 450 µl of DMSO to a rough concentration of 20 mM. These 86 peptides were then tested as crude mixtures for binding towards VBC by two orthogonal techniques, DSF and BLI.

4.2.3.1 Differential Scanning Fluorimetry

The DSF assay, also known as thermal shift, measures the thermal stability of a protein and subsequent change in its melting temperature (T_M) upon binding of a ligand. The denaturation curve of VBC in buffer (Figure 4.5) reported an unfolding event at 51 °C, whereas in the presence of a nanomolar binder, HIF-1α 19-mer peptide, used as positive control, the melting temperature of the complex is increased to 59 °C, as expected. The bicyclic peptides were tested at a final concentration of 500 µM, whilst the protein concentration was kept at 7.5 µM. The analysis of the thermal denaturation curves (Figure 4.5) revealed that most of the peptides exhibited a level of fluorescence close to the detection limit at the lowest temperature and, therefore, no increase in fluorescence upon unfolding of the protein could be

readily detected. This could potentially be explained by the fact that the bicyclic peptides can, on their own, mimic protein interfaces with exposed hydrophobic patches that can bind SYPRO Orange. This phenomenon makes the assay unreliable for these peptides hence the results must be taken with caution. In the cases where it was possible to observe an appropriate denaturation curve, negative shifts in the melting temperature were generally observed, suggesting destabilisation of the VBC complex in the presence of the crude bicyclic peptide mixture.

All together, the thermal shift experiment results were inconclusive.

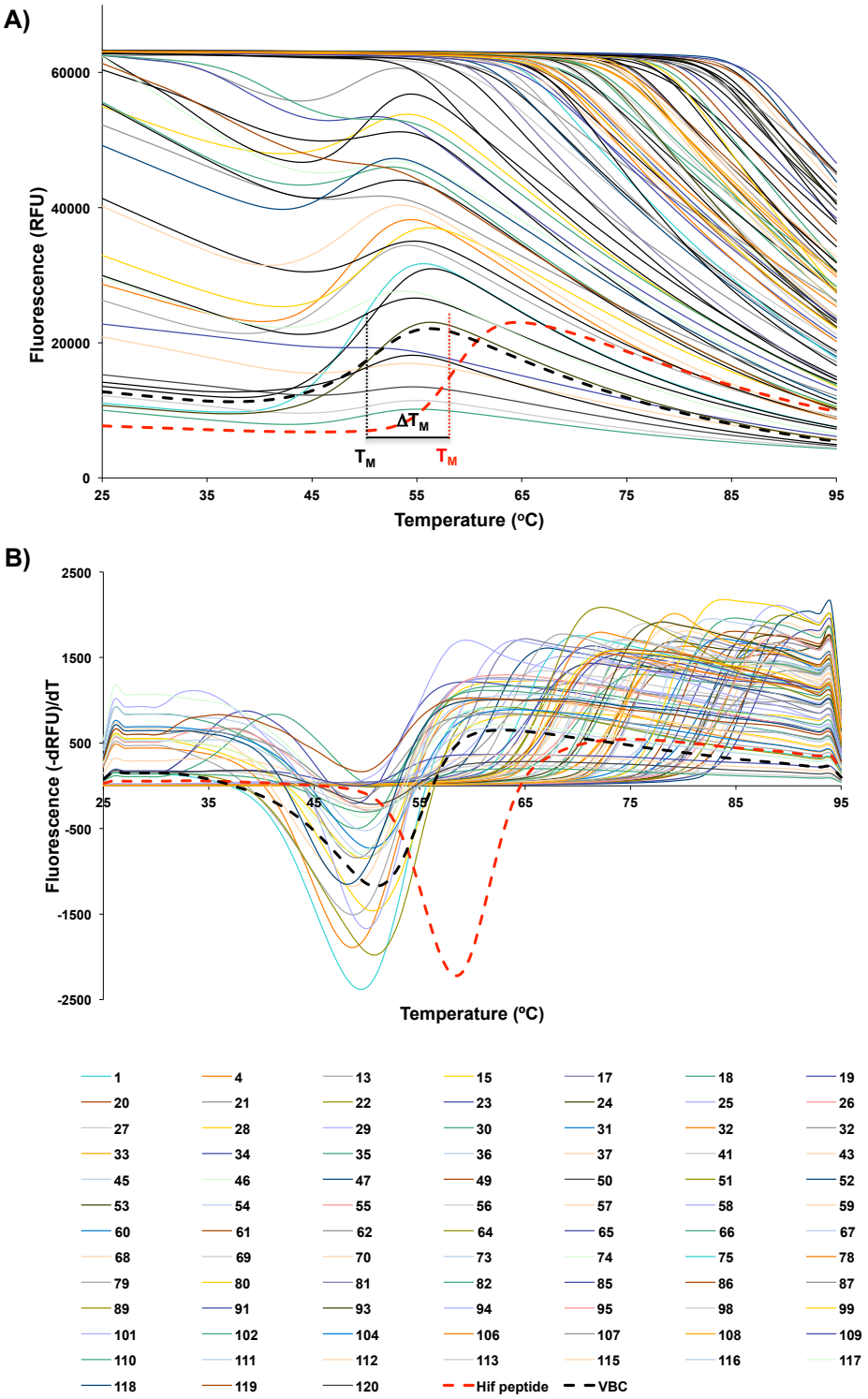


Figure 4.5 – Differential scanning fluorimetry experiment. A) Thermal denaturation curves of VBC (7.5 μM) in the presence of 500 μM of the each of the 86 bicyclic peptides tested. **B)** Derivatives of RFU relative to temperature are plotted for VBC (7.5 μM) in the presence of 500 μM of the each of the 86 bicyclic peptides tested. The melting temperatures of VBC and VBC in the presence of the HIF-1α 19-mer peptide (VBCH) are marked in black and red, respectively. The assay was performed in 100 mM Bis-tris propane pH 7.5, 1 mM DTT, 10% DMSO, 0.02% Tween-20.

4.2.3.2 Biolayer interferometry

The attention was then turned to the characterisation of peptide binding using BLI. The experiment was performed stepwise. Initially, the 86 bicyclic peptides were tested in a single-point concentration assay and then, dose-response curves were obtained from the ones originating a measurable response in the first round, to assess concentration-dependent effects.

4.2.3.2.1 Single-point concentration assay

In this assay, streptavidin-coated biosensor tips were loaded with biotinylated VBC or VBCH and were dipped into a solution of bicyclic peptide (approximately 500 μM). The assay was designed with an association step of 60 seconds (when the tip is dipped into the well containing the ligand) and a dissociation step of 180 seconds (when the tip is dipped in a well containing only buffer and the ligand dissociates from the protein). All the 86 crude mixtures of cyclised peptides were tested in this assay against VBC and VBCH. Of these, 17 peptides showed a positive response against VBC and 41 peptides showed positive response against VBCH (15 being common to both assays) – data in Appendix I. This yielded 43 unique bicyclic peptides, which were taken forward to a concentration-dependent response assay against VBC and VBCH.

4.2.3.2.2 Concentration-dependent response assays

The aim of this experiment was to validate and further characterise the peptides previously identified in the single point screening as actual binders. To this end, the peptides (still as crude mixture) were titrated against VBC and VBCH starting at a top concentration of approximately 500 μM and using three-fold dilutions to a minimum concentration of approximately 2 μM. The set

up of the experiment was maintained. In some cases, the data at the top concentration of the peptides was consistent with aggregation, thus the measurements at this concentration were excluded from the analyses. Of the 43 bicyclic peptides tested, 19 showed concentration-dependent responses against VBC and VBCH with estimated K_d values in the range of double-digit micromolar (Table 4.7, sensorgrams in Appendix I).

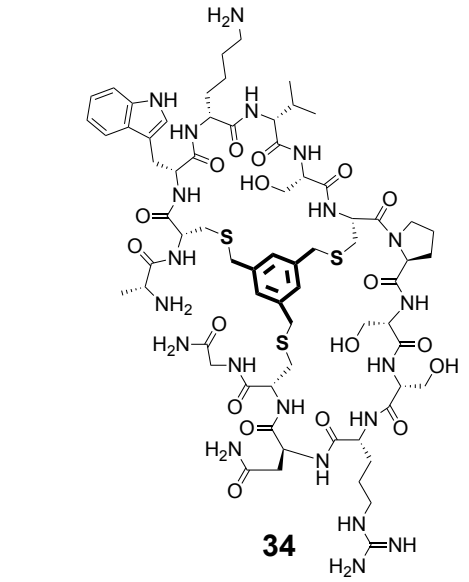
Table 4.7 – Concentration-dependent response BLI assays. Bicyclic peptides (crude mixture) were titrated against VBC and VBCH and 19 peptides showed concentration-dependent response. K_d values were determined based on the fitting associated to the response obtained in function of the concentration using a 1:1 model. The R^2 values represent how well the data is fitted to the model. The assays were conducted in 100 mM Bis-tris propane pH 7.5, 1 mM DTT, 10% DMSO, 0.02% Tween-20.

	VBC		VBCH	
Peptide	**K_d (μM)**	**R^2**	**K_d (μM)**	**R^2**
15	33	0.9913	72	0.9969
22	26	0.9935	69	0.9995
24	12	0.9972	26	0.9536
27	8	0.9855	29	0.9769
30	8	0.9902	8.8	0.9319
34	19	0.8274	41	0.9708
37	2.7	0.9962	4.1	0.9690
41	6.7	0.9317	13	0.9750
64	44	0.9943	44	0.9939
66	37	0.9868	18	0.9492
73	51	0.9772	60	0.9936
74	180	0.9986	-	-
86	11	0.8741	59	0.9955
93	9.2	0.8584	81	0.9989
107	100	0.9967	-	-
108	21	0.9912	65	0.9878
109	35	0.9859	55	0.9997
113	14	0.9345	48	0.9960
115	23	0.9948	49	0.9995

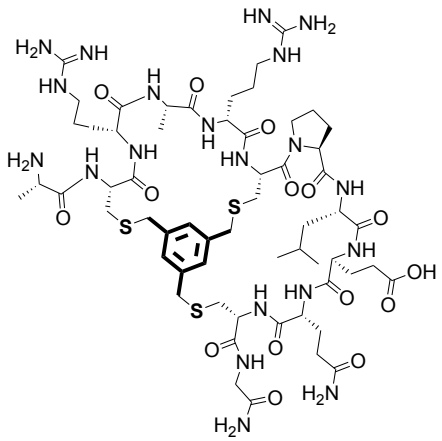
These 19 bicyclic peptides were selected for scale-up synthesis, purification and in-depth characterisation of binding towards VBC.

4.2.4 Large-scale synthesis and purification of the hit peptides

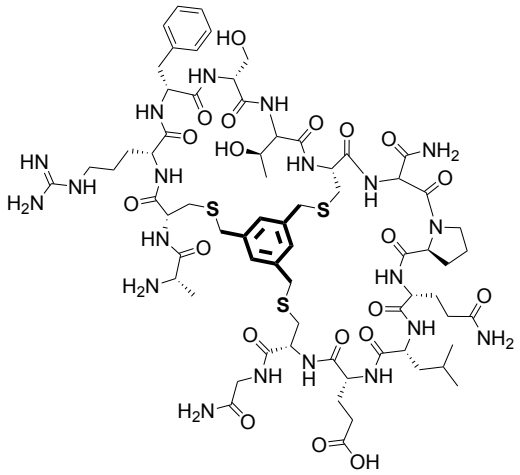
The 19 bicyclic peptides selected were synthesised in a larger scale. The peptide synthesis and cyclisation of the peptides was performed as described before. The following step, HPLC purification, was hindered as the solubilisation of the peptides proved a difficult task. Different ratios of H_2O:ACN were tested in combination with variable percentages of DMSO. Also the presence (< 0.1%) of TFA, formic acid or ammonium sulfate in the solutions was tested. In addition, the peptide solutions were sonicated in an attempt to improve the solubilisation. Nonetheless, some peptides were not soluble in any of the conditions tested and due to these solubility issues the HPLC purification was only carried out for 12 of these peptides. The purification was performed in acidic conditions and the gradient of H_2O:ACN was optimised for each peptide. However, the yield of the purifications was very poor (10-20%) and only a small amount of purified bicyclic peptide was recovered (1-3 mg). This low yield could be explained by two factors related to the conditions required for the solubilisation of the peptides that, on one side could have prevented the peptides to be retained in the HPLC column and, on the other side, could have prevented the detection by the UV-Vis detector due to the low concentration and, therefore, prevented their recovery. According to the analytical HPLC-MS traces (shown in Chapter 6), of the 12 recovered peptides only peptides **34**, **66**, **86**, **109**, **113** and **115** (Figure 4.6) were satisfactorily pure (> 95%) to be tested biophysically for binding towards VBC. With the exception of peptide **66,** which was identified 5 times in the screening of VL1 against VBC, the remaining peptides were observed only as a single instance during the phage display screening.



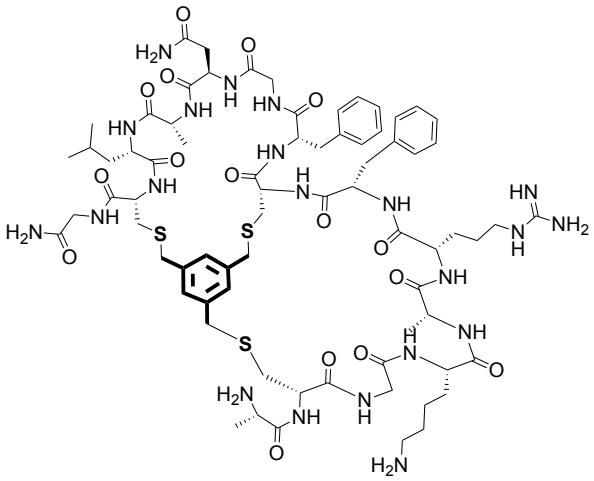
34



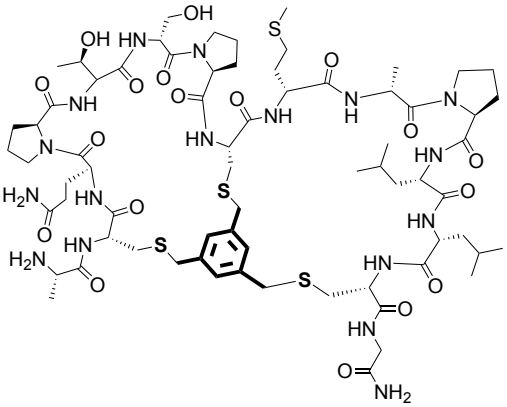
66



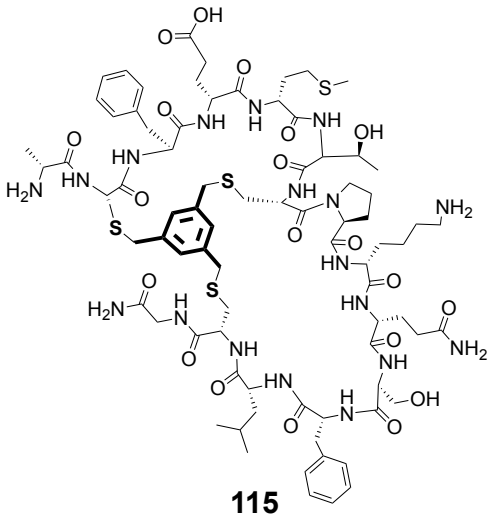
86



109



113



115

Figure 4.6 – Bicyclic peptides for validation. Chemical structures of bicyclic peptides **34**, **66**, **86**, **109**, **113** and **115**.

4.2.5 Validation of binders

The binding of the bicyclic peptides **34**, **66**, **86**, **109**, **113** and **115** to VBC was investigated by three orthogonal biophysical techniques – BLI, DSF and ligand-observed NMR. The thermal shift experiment results determined neglegible shifts (< 1 °C) in the melting temperature of VBC in the presence of the bicyclic peptide (Figure 4.7). Moreover, in the cases of peptides **86** and **109** these shifts were negative, suggesting a destabilisation of the complex. The extent of the shifts observed were in any cases minimal, suggesting weak binding if at all.

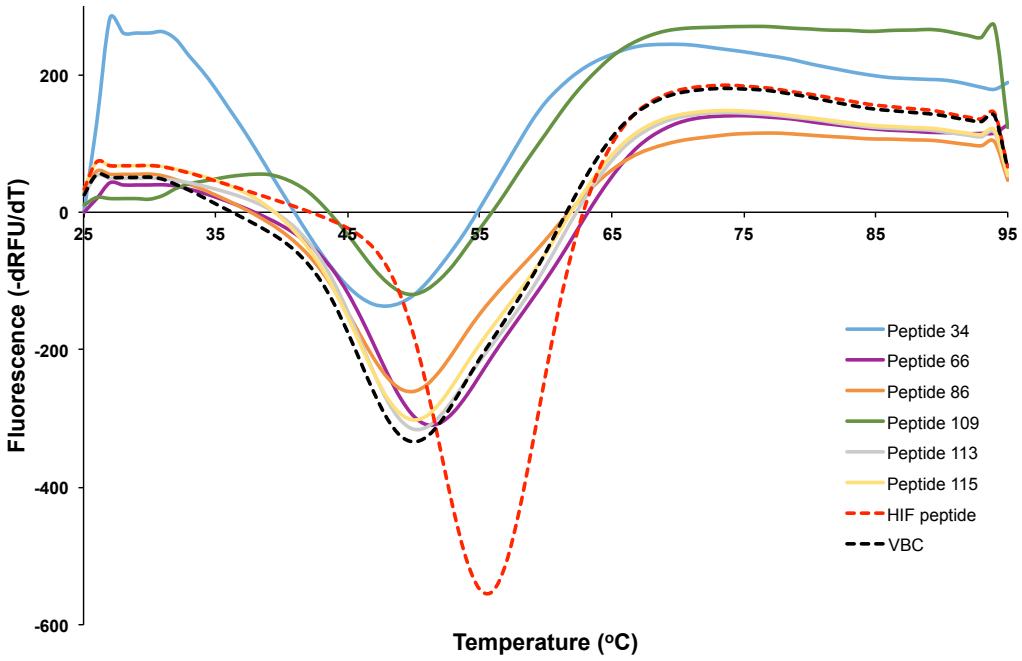


Figure 4.7 – Derivative curves from the DSF experiment. Bicyclic peptides (250 μM) were mixed with 7.5 μM VBC in 100 mM Bis-tris propane pH 7.5, 1 mM DTT, 10% DMSO, 2.5x SYPRO Orange and the denaturation curves measured. The curves shown are the average of a duplicate. For reference, the curves of VBC (black dashed) and VBC in the presence of HIF-1α 19mer (red dashed) are shown.

Next, the peptides were titrated by BLI against VBC. The range of concentrations tested varied from 500 μM to 2 μM. It was observed that after the purification only bicyclic peptides **34**, **66** and **109** showed

concentration-dependent response by BLI (Figure 4.8). However, the responses observed for peptides **34** and **66** showed systematic deviations between the experimental data and the fitted values, suggesting that the simple 1:1 binding model used did not accurately reproduce the data. Thus, other effects in addition to genuine and reversible 1:1 binding were likely to be accounting for the observed signal.

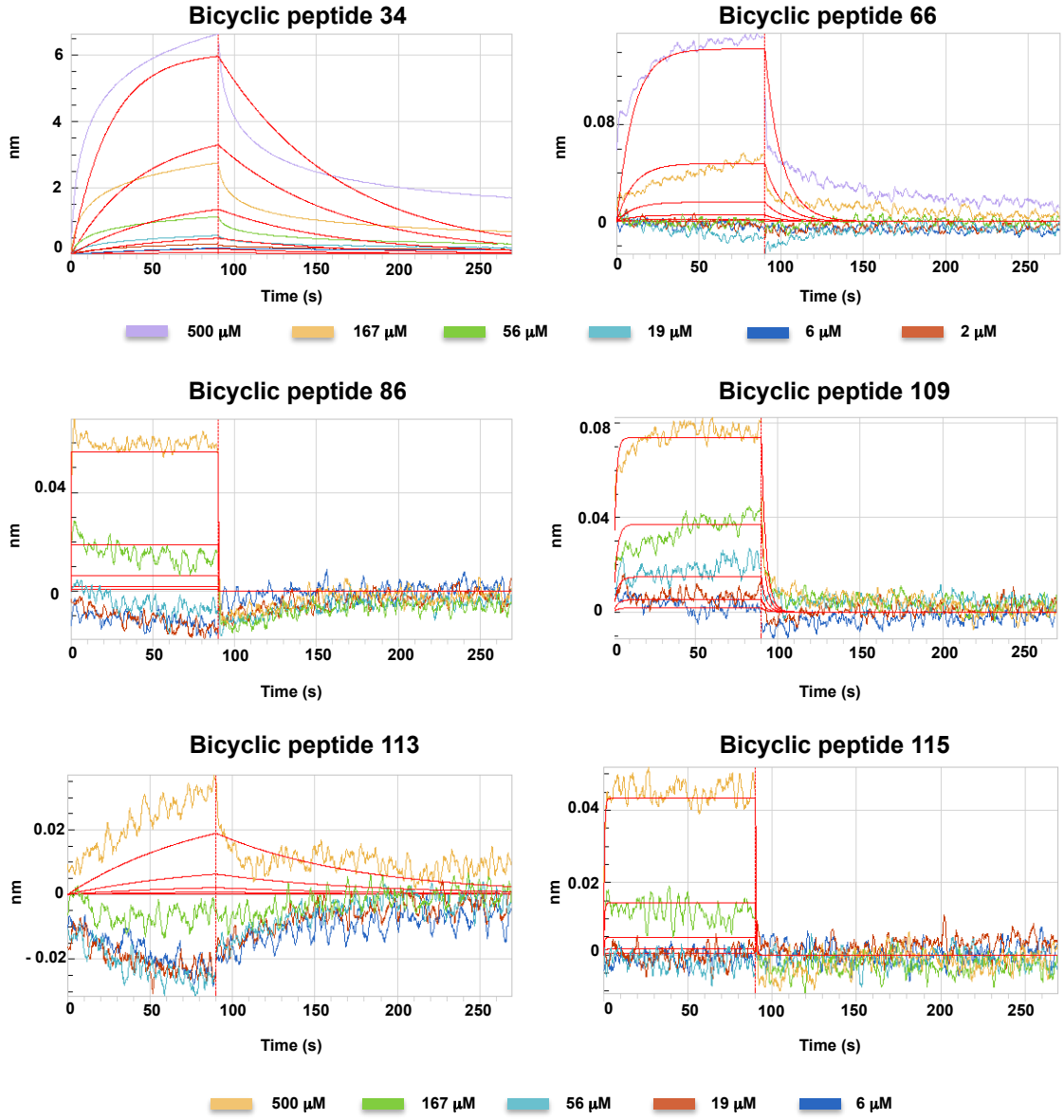


Figure 4.8 – BLI experiments. Bicyclic peptides were titrated against VBC in a concentration range from 500 µM – 2 µM. Biotinylated VBC was immobilised in streptavidin-coated sensors. The assays were conducted in 100 mM Bis-tris propane pH 7.5, 1 mM DTT, 10% DMSO, 0.02% Tween-20.

Finally, peptides **34**, **66**, **86** and **109** were also tested by ligand-observed NMR (the amount left of peptides **113** and **115** was not enough to test them in this experiment) - Figure 4.9. In the CPMG experiment, the ^{1}H NMR spectrum of the sample was recorded in the absence and in the presence of VBC and, upon a binding event, it was expected to observe a decrease in the intensity of the peaks. Peptide **34** showed no binding, peptides **66** and **86** suggested a weak interaction and peptide **109** showed moderate binding in this experiment.

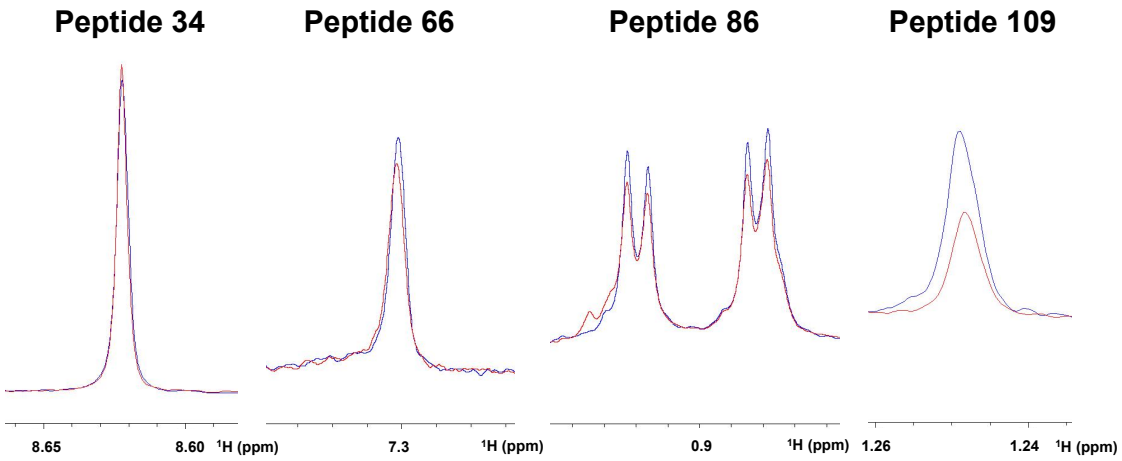


Figure 4.9 – Ligand-observed NMR experiments. CPMG experiments were performed to assess binding of the bicyclic peptides to VBC and isolated peaks from each experiment are shown. Blue corresponds to the ^{1}H spectrum of 1 mM of peptide and red corresponds to the ^{1}H spectrum of 1 mM of peptide in the presence of 30 μM of VBC.

The combination of results obtained from the different techniques (Table 4.8) suggested these bicyclic peptides to be weak binders of the VBC complex. Efforts were conducted to scale up the synthesis and purification of these peptides in order to perform a greater biophysical evaluation and use other techniques such as, isothermal titration calorimetry, X-ray crystallography and protein-observed NMR. However, their lack of solubility constituted the major obstacle for the cyclisation and purification steps and prevented the continuation of the work.

Table 4.8 – Validation of bicyclic peptides as binders. The peptides were tested by DSF, BLI and NMR. ND means not determined; X means no binding observed and a tick means binding observed.

Peptide number	Peptide sequence	Peptide format	ΔT_M (°C)	BLI K_D (M)	NMR (CPMG)
34	A**C**WKVS**C**PSSRN**C**G	4x5	ND	9.8×10^{-4}	X
66	A**C**RAR**C**PLEQ**C**G	3x4	+ 0.4	(bad fitting)	✓
86	A**C**RFST**C**NPQLE**C**G	4x5	- 0.3	X	✓
109	A**C**GKARF**C**FGNAL**C**G	5x5	- 0.7	9.4×10^{-4}	✓
113	A**C**QPTSP**C**MAPLL**C**G	5x5	+ 0.8	X	ND
115	A**C**FEMT**C**PKQSFL**C**G	4x6	+ 0.3	X	ND

4.3 Discussion

Protein-protein interactions are challenging targets due to their large interface areas and relatively flat and featureless topologies. One avenue to overcome this challenge is by using mimetics of protein surfaces. Bicyclic peptides arise in this context as ideal candidates to accomplish this goal. They provide a large surface area with potential for tight binding affinities and selectivity and they present a constrained arrangement, which might enhance the binding properties as well as reduce the entropic penalty upon binding.[63]

The phage display screening of bicyclic peptides performed against VBC and VBCH generated a set of 126 peptides with different formats and sequences. Libraries VL1 and VL3 yielded no unique consensus sequences, although the screening of VL3 against VBC originated one highly enriched sequence. Conversely, the screening of library VL2 resulted in some convergence, with the same motifs being repeated in different peptides. Moreover, five sequences from the VL2 library were simultaneously identified in the two independent screenings against VBC and VBCH, hinting at some specific binding. In addition, in the last round of biopanning, the 4x5 and 5x4 format of bicyclic peptides seemed to be preferred for binding to VBC. In fact, the same trend has been reported against a different target,[143] suggesting that the VL2 library might contain the preferred format of bicyclic peptides to act as binders. Taking into account the lack of unique consensus obtained from the phage display, all the bicyclic peptides identified were tested for binding. From the dose–response BLI experiments, 19 peptides showed binding affinities in the micromolar range, which was encouraging. The fact that the peptides had been tested as crude mixtures up until this point could be masking tighter binding affinities or hiding false positives. For this reason, some of these

peptides (**34, 66, 86, 109, 113** and **115**) were tested again in a purified state, which led to their confirmation as very weak binders. None of the peptides showed binding in all the three biophysical experiments performed (DSF, BLI and NMR), the K_d value determined by BLI for peptide **109** was close to the millimolar range (940 μM), and the BLI response for peptide **34** was unexpectedly high with the curves not reaching a plateau, thus suggesting aggregation. Despite weak, if these bicyclic peptides were real binders they could provide starting points for their optimisation into potent ligands. This could be achieved by small modifications at the side chain level or even replacement of some residues by D-amino acids,[154] which could be guided by X-ray crystallography and protein-observed NMR, for instance. Having this in mind, it was decided to scale up the synthesis and purification of the bicyclic peptides tested. However, this proved very challenging as the peptides where extremely insoluble, complicating the purification step, where, invariably, large amounts of peptide were lost. Despite much effort, which involved different solvents, different approaches of solubilisation and even the addition of arginine residues at the N-terminus of the sequence of the bicyclic peptides to make them more hydrophilic, this was not achieved.

Notwithstanding the solubility concerns, from the phage display screening it was expected to obtain reasonably potent (low micromolar)[155] and selective hits, which was not the case when the peptides were tested after being purified. The results obtained suggested that the phage display screening was not successful with the two targets tested, VBC and VBCH. It was reasoned that, in this case, the failure to obtain potent ligands could have happened because: 1) VBC is not ligandable by bicyclic peptides; 2) the immobilisation of the target on the magnetic beads could have led to the

hindering of potential interaction surfaces; or 3) failure of the technology due to inherent challenges and limitations of solubility. It is noteworthy that much success has been achieved with symmetrical bicyclic peptides (namely the 6x6 library)[148,156-159] but not much has been reported with the VL1, VL2 and VL3 libraries. It is therefore proposed that it could be of interest in the future to explore the 6x6 library against these targets instead.

Phage display of bicyclic peptides is a powerful approach for the identification of ligands of protein surfaces. Nonetheless, as any other technique, it has its drawbacks and might not be applicable to all protein targets. In this case, it was also applied, to our knowledge for the first time, to a multi-subunit protein complex. Bicyclic peptides, as potential mimetics of protein surfaces, can be very challenging molecules to work with due to their observed low solubility. In fact, chemical strategies to improve the solubility of bicyclic peptides and their biological activity are an active area of research. To this end, recent efforts have led to the design of more polar and less hydrophobic core scaffolds compared to TBMB, for example, 2,4,6-tris(bromomethyl)-s-triazine, TBMT, and 1,1',1"-(1,3,5-triazinane-1,3,5-triyl)tris(2-bromoethanone), TATB. TBMT and TATB are structurally very similar to TBMB but present very different chemical properties.[160]

In summary, there is no universal technique to identify binders of a certain protein surface and each target, approach and technology must be studied as a particular case.

CHAPTER 5

CONCLUDING REMARKS

5.1 Concluding remarks

Cullin RING E3 ubiquitin ligases play important roles in cellular homeostasis by controlling the abundance of different regulatory proteins. In many diseases including cancer this process is deregulated, which motivates the search for chemical probes targeting these complexes.[161,162] Targeting CRLs for drug development is, consequently, an ambitious goal of mounting interest. Furthermore, the emerging role of CRLs as targets for induced protein degradation by glue molecules[163] and PROTACs[164] emphasise their potential as therapeutic targets.[25,165] However, not all protein targets are druggable and assessing their ability to bind small molecules (ligandability) is critical. This project aimed at exploring the ligandability of a representative multi-subunit CRL complex, the $CRL2^{VHL}$, combining three different avenues and a multiplicity of biochemical, biophysical and structure-based design tools.

Initially, a structural and biophysical study was conducted, resulting in the first crystal structure of the whole $CRL2^{VHL}$ complex, in the biophysical characterisation of a particular PPI within the complex (Cul2–VBC), and in the identification of amino acids responsible for the selective recruitment of Cul2 over Cul5 in EloBC-containing CRL complexes. This new structure revealed details of Cullin flexibility and disclosed a new conformation for the RING domain of Rbx1 that resembles a pose on a trajectory from the inactive to the active state of the enzyme. The challenges associated with the determination of the crystal structure were a consequence of the low-resolution of the data collected. Multi-subunit large protein complexes are usually very dynamic, thus hard to crystallise and to produce high-resolution data, which constitutes a problem for X-ray crystallography.[71] This limitation can now be overcome by cryo-electron microscopy, a rising structural biology technique that allows

structure determination of proteins in their native environment, with no need for crystallisation of the sample. Initially, this technique was only applicable to complexes on the megadalton size range, but recent advances of the technology have now permitted the determination of enzyme structures as small as 170 kDa.[166] In the future, this could be a potential way of exploring structurally, the dynamic features of the $CRL2^{VHL}$ complex. In addition, small-angle X-ray scattering (SAXS), a biophysical technique that allows the determination of the overall shape and structural transitions of biological molecules in solution[167] could be useful to investigate the inherent flexibility of the pVHL β-domain in the context of the whole complex.

The second avenue pursued involved the structure-guided design of peptide molecules to probe the Cul2–VBC interaction in the $CRL2^{VHL}$ complex. Linear and cyclic peptides, short and long sequences were tested, however, targeting this interaction with peptides proved very challenging. The work led to the identification of some ligands that, nonetheless, exhibited weak binding affinities and could be optimised further in the future. The Cul2–VBC interaction appears to be a tertiary protein-protein interaction,[51] involving multiple epitopes. In fact, as seen in the crystal structure, the interaction is directed by three points of contact between pVHL, Cul2 and EloC. Despite the N-terminal loop of Cul2 being important for the specificity of the interaction, by itself it could not drive a tight binding event. Therefore, peptides that target only one of these interaction sites would likely not be able to mimic the native interaction and thus would not block the PPI site effectively. Considering these results combined with the merely hydrophobic nature of the EloC site, it can be concluded that this is a very low ligandability binding site.

Lastly, the surface of VBC was probed with bicyclic peptides originated from three different phage display libraries in an unbiased approach. There were concerns about the outcome of the phage display screening, as no unique consensus sequence was observed. Nevertheless, once proved to bind VBC, the diverse peptide sequences identified could provide starting points for ligand optimisation. The biophysical characterisation of these peptides binding to VBC, however, proved very problematic due to the lack of solubility of these molecules and it has remained inconclusive. Based on this experience, it seemed crucial that the next step in optimising this exciting new technology would be the enhancement of the solubility and physicochemical properties of the bicyclic peptides.

All together, the work developed and presented in this thesis contributes to a better understanding of the structure and functioning of the $CRL2^{VHL}$ and provides important insights into the tractability of this complex. Moving forward, it is expected that the information contributed by this work could facilitate the design of chemical tools and modulators of the VHL ligase to study CRLs biology.

CHAPTER 6

EXPERIMENTAL

6.1 Molecular biology methods

6.1.1 Plasmids

The plasmid encoding his-tagged pVHL short isoform (54-213) was cloned by Dr. Inge van Molle, a former post-doctoral researcher in the Ciulli laboratory, and the EloBC plasmid was a gift from Dr. Alex Bullock (SGC, Oxford). The truncated Cullin-2 NTD plasmids were generated from an original plasmid kindly offered by Dr. Alexander Buchberger (University of Würzburg). The plasmid for the co-expression of full-length Rbx1 and Cul2 with a N-terminal Dac tag (affinity tag based on penicillin-binding protein 5)[89] was a gift from Dr. Mark Peggie (MRC-PPU/DSTT). All the plasmids and respective mutations are summarised in Table 6.1.

Table 6.1 – Summary of all the constructs.

Plasmid ID	Protein	Vector	Resistance	Tags	Cleavage site	Expected Mass (Da)	Measured Mass (Da)
pIVM02	$pVHL_{54-113}$	pHAT4	Ampicillin	6xHis	TEV	18676	18676
TFC024	$pVHL_{54-113}$ V181G	pHAT4	Ampicillin	6xHis	TEV	18634	18635
TFC025	$pVHL_{54-113}$ D187K	pHAT4	Ampicillin	6xHis	TEV	18689	18690
TFC026	$pVHL_{54-113}$ K159E	pHAT4	Ampicillin	6xHis	TEV	18677	18678
pIVM26	$EloB_{1-104}/EloC_{17-112}$	pCDF1b-DUET	Streptomycin	-	-	EloB:11733 EloC: 10964	EloB:11733 EloC: 10964
TFC027	$EloB_{1-104}/EloC_{17-112}$ M105A	pCDF1b-DUET	Streptomycin	-	-	EloB:11733 EloC: 10904	EloB:11733 EloC: 10904
TFC028	$EloB_{1-104}/EloC_{17-112}$ F109A	pCDF1b-DUET	Streptomycin	-	-	EloB:11733 EloC: 10888	EloB:11733 EloC: 10888
TFC012	$Cul2_{1-163}$	pTvHR21-SGC	Ampicillin	MBP-6xHis	Fxa/TEV	64211	
TFC014	$Cul2_{1-163}$	pTvHR21-SGC	Ampicillin	SUMO-6xHis	TEV/SENP1	33220	
TFC017	$Cul2_{1-384}$	pTvHR21-SGC	Ampicillin	MBP-6xHis	Fxa/TEV	89968	
TFC019	$Cul2_{1-384}$	pTvHR21-SGC	Ampicillin	SUMO-6xHis	TEV/SENP1	58977	
TFC020	$Cul2_{1-384}$ V340R, L344D, V384R	pTvHR21-SGC	Ampicillin	MBP-6xHis	Fxa/TEV	90084	
TFC021	$Cul2_{1-384}$ V340R, L344D, V384R	pTvHR21-SGC	Ampicillin	MBP-6xHis	PreScission/TEV	90513	
TFC003	Cullin-2/Rbx1	pFastBacDUAL	Ampicillin	Dac	TEV	Cul2: 87184 Rbx1:12274	Cul2: 87184 Rbx1:12185
TFC029	Cullin-2 K114E/Rbx1	pFastBacDUAL	Ampicillin	Dac	TEV	Cul2: 87185 Rbx1:12274	Cul2: 87159 Rbx1:12185
TFC030	Cullin-2 Q111L/Rbx1	pFastBacDUAL	Ampicillin	Dac	TEV	Cul2: 87169 Rbx1:12274	Cul2: 87172 Rbx1:12185
TFC031	Cullin-2 F39A/Rbx1	pFastBacDUAL	Ampicillin	Dac	TEV	Cul2: 87108 Rbx1:12274	Cul2: 87111 Rbx1:12185
TFC032	Cullin-2 Y43A/Rbx1	pFastBacDUAL	Ampicillin	Dac	TEV	Cul2: 87092 Rbx1:12274	Cul2: 87097 Rbx1:12185
TFC033	Cullin-2 V47A/Rbx1	pFastBacDUAL	Ampicillin	Dac	TEV	Cul2: 87156 Rbx1:12274	Cul2: 87159 Rbx1:12185
TFC034	Cullin-2 P5A/Rbx1	pFastBacDUAL	Ampicillin	Dac	TEV	Cul2: 87158 Rbx1:12274	Cul2: 87161 Rbx1:12185
TFC035	Cullin-2 R6A/Rbx1	pFastBacDUAL	Ampicillin	Dac	TEV	Cul2: 87099 Rbx1:12274	Cul2: 87099 Rbx1:12185
TFC036	Cullin-2 L3A/Rbx1	pFastBacDUAL	Ampicillin	Dac	TEV	Cul2: 87142 Rbx1:12274	Cul2: 87147 Rbx1:12185
TFC038	Cullin-2 L3G/Rbx1	pFastBacDUAL	Ampicillin	Dac	TEV	Cul2: 87128 Rbx1:12274	Cul2: 87133 Rbx1:12185
TFC039	Cullin-2 N36A/Rbx1	pFastBacDUAL	Ampicillin	Dac	TEV	Cul2: 87141 Rbx1:12274	Cul2: 87144 Rbx1:12185
TFC054	$pVHL_{54-113}$ K159Q D187Y S183R	pHAT4	Ampicillin	6xHis	TEV	18794	18794
TFC055	$SOCS2_{32-198}$ Q164K Y190D R186S	pLIC-SGC1	Ampicillin	6xHis	TEV	19128	19128

The Δ = 89 Da observed for Rbx1 is consistent with the removal of initiator methionine from protein N-terminus, then acetylation of the new N-terminus.

6.1.2 Cloning of the truncated Cul2 NTD constructs

All Cul2 constructs were obtained using pTvHR21-SGC, a plasmid that contains a C-terminal 6xHis tag (Figure 6.1). Initially MBP-TEVsite or SUMO tags were amplified by PCR from their parental plasmids using specific primers (Table 6.2) that allowed the insertion of EcoRI, NdeI and BamHI restriction sites at the 5' end and XhoI restriction site at the 3' end. Using NdeI and XhoI restriction enzymes, the amplified inserts containing MBP-TEVsite-6xHis or SUMO-6xHis tags, were cloned into pTvHR21-SGC using the DNA Ligation kit (DNA Ligation kit, TaKaRa), following the manufacturer guidelines.

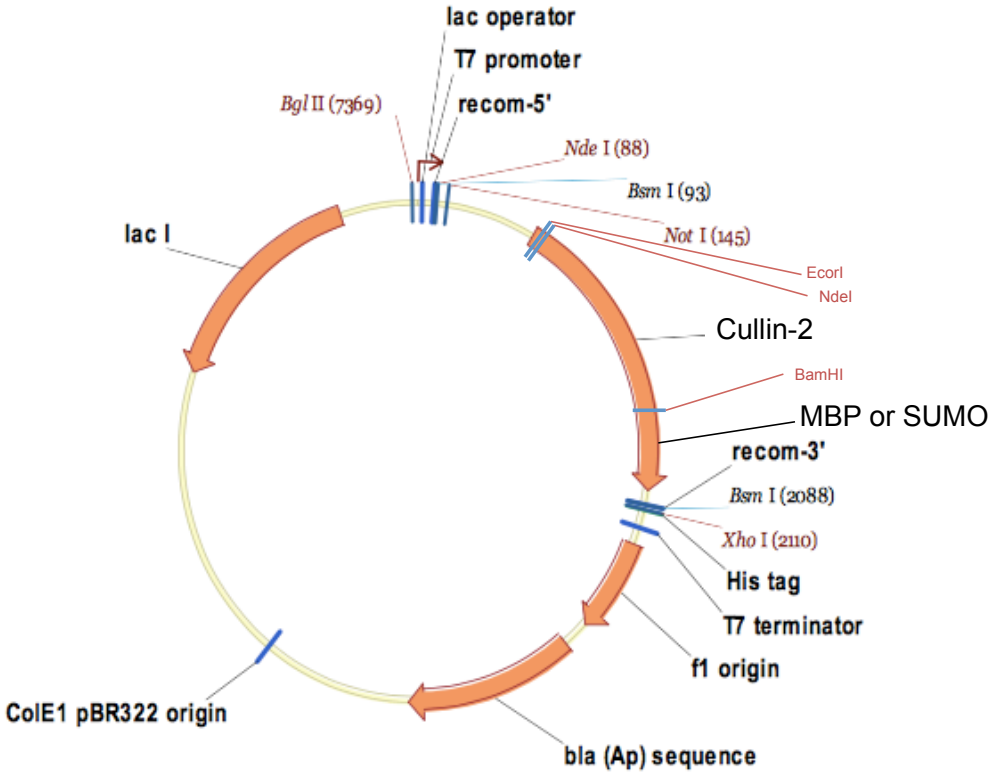


Figure 6.1 – Vector map representative of the final Cullin-2 plasmid.

Next, Cullin-2 inserts of different lengths (1-163 or 1-384) were amplified by PCR using specific primers (Table 6.2) that introduced different cleavage sites and restriction sites (EcoRI and NdeI at the 5' end and BamHI at the

3’ end). The inserts were digested with NdeI and BamHI and ligated to the respective MBP- or SUMO-containing plasmid using the same protocol as described before. The identity of the constructs was verified by DNA sequencing.

Because the DNA sequence of full length Cullin-2 contained two NdeI restriction sites, before starting the cloning these sites were mutated by site-directed mutagenesis, using the primers reported in Table 6.2.

Table 6.2 – Primers used for the cloning of the Cul2 constructs.

Primer name	**Primer sequence 5' --> 3'**
Cul2 SDM 1 Fwd	GTGTCCACTGGTTTACCTCACATGATTCAGGAGCTGC
Cul2 SDM1 Rev	GCAGCTCCTGAATCATGTGAGGTAAACCAGTGGACAC
Cul 2 SDM 2 Fwd	CTATTTGGGCAAACCACATGTAGCCATGGTTACAAC
Cul2 SDM2 Rev	GTTGTAACCATGGCTACATGTGGTTTGCCCAAATAG
EcorI-NdeI-BamHI-SUMO (Fwd)	GAATTCCATATGGACTCGGATCCGGCAGCGGCAGCGGCTCT
SUMO-XhoI (Rev)	GGCAGCCTCGAGGCTGCCACCACCGGTTTGTTCCTG
EcoRI-NdeI-BamHI-MBP (Fwd)	GAATTCCATATGGACTCGGATCCATGAAGATTGAAGAAGGC
MBP-TEVsite-XhoI (Rev)	GGCTGCCTCGAGCGAGCCACCCTGGAAGTACAGGTT
EcoRI-NdeI-Cul2 (Fwd)	GAATTCCATATGTCTTTGAAACCAAGAGTA
Cul2-BamHI 1-163 (Rev)	GGATCCCAGCATTCGGATAAGGATG
Cul2-BamHI 1-384 (Rev)	GGATCCAACAGACTTAGGTTCTCTG

6.1.3 Site-directed mutagenesis

The pVHL, SOCS2, EloC and Cul2 mutants were prepared by PCR-based method using the respective expression vectors encoding for the wild type proteins as template. The amplification of the expression vectors was performed using the KOD hot-start DNA polymerase (EMD Millipore), following the manufacturer guidelines and specific pairs of primers (Table 6.3) for the introduction of the desired mutation. The PCR products were treated with DpnI

enzyme (NEB) to digest the parental DNA and then transformed in DH5α *E. coli* cells. The mutations were confirmed by DNA sequencing.

Table 6.3 – Primers used for the mutations of Cullin-2, pVHL, EloC and SOCS2.

Protein	Mutation	Forward primer 5' -> 3'	Reverse primer 5'->3'
Cullin-2	L3A	CAGGGCGGATCCATGTCTGCGAAACCAAGAGTAGTAG	CTACTACTCTTGGTTTCGCAGACATGGATCCGCCCTG
	L3G	GTTCCGCGTGGATCCATGTCTGGGAAACCAAGAGTAGTAG	CTACTACTCTTGGTTTCCCAGACATGGATCCACGCGGAAC
	P5A	CCGCGTGGATCCATGTCTTTGAAAGCAAGAGTAGTAGATTTTGATGAAACATGG	CCATGTTTCATCAAAATCTACTACTCTTGCTTTCTCAGACATGGATCCACGCGG
	N36A	GGAATACGTCGAAAGAGCAACATGGGCTGACCGTTTCTCAGATATCTATGCTTTATG	CATAAAGCATAGATATCTGAGAAACGGTCAGCCCATGTTGCTCTTTCGACGTATTCC
	F39A	CGTCGAAAGAGCAACATGGAATGACCGTGCCTCAGATATCTATGCTTTATGTGTGGCC	GGCCACACATAAAGCATAGATATCTGAGGCACGGTCATTCCATGTTGCTCTTTCGACG
	Y43A	GGAATGACCGTTTCTCAGATATCGCTGCTTTATGTGTGGCCTATCCTGAACCCC	GGGGTTCAGGATAGGCCACACATAAAGCAGCGATATCTGAGAAACGGTCATTCC
	V47A	CAGATATCTATGCTTTATGTGCGGCCTATCCTGAACCCCTTGGAG	CTCCAAGGGGTTCAGGATAGGCCGCACATAAAGCATAGATATCTG
	Q111L	GGTATCTCAACACCCTGTTTATTAAAAAGAATAAATTAACAGAAGCGGACC	GGTCCGCTTCTGTTAATTTATTCTTTTTAATAAACAGGGTGTTGAGATACC
	K114E	GTATCTCAACACCCAGTTTATTGAAAAGAATAAATTAACAGAAG	CTTCTGTTAATTTATTCTTTTCAATAAACTGGGTGTTGAGATAC
pVHL	K159E	CACTGCCAGTGTATACTCTGGAAGAGCGATGCCTCCAGGTTGTCC	GGACAACCTGGAGGCATCGCTCTTCCAGAGTATACACTGGCAGTG
	K159Q	CCAATTGAGATCCCTAACAATCTAAGGCGTTCGCAGCCATCAGCAGTTCC	CCAATTGAGATCCCTAACAATCTAAGGCGTTCGCAGCCATCAGCAGTTCC
	V181G	GAATTACAGGAGACTGGACATCGGCAGGTCGCTCTACGAAGATC	GATCTTCGTAGAGCGACCTGCCGATGTCCAGTCTCCTGTAATTC
	S183R+D187Y	CCAATTGAGATCCCTAACAATCTAAGGCGTTCGCAGCCATCAGCAGTTCC	CGTCAGGCGGCTCTACGAATATCTGGAAGACCACCCAAATGTGC
	D187K	CATCGTCAGGTCGCTCTACGAAAAACTGGAAGACCACCCAAATG	CATTTGGGTGGTCTTCCAGTTTTTCGTAGAGCGACCTGACGATG
EloC	M105A	GAAATTGCACTGGAACTGCTGGCGGCTGCGAACTTCTTAGATTGTTAG	CTAACAATCTAAGAAGTTCGCAGCCGCCAGCAGTTCCAGTGCAATTTC
	F109A	GGAACTGCTGATGGCTGCGAACGCCTTAGATTGTTAGGGATCTCAATTGG	CCAATTGAGATCCCTAACAATCTAAGGCGTTCGCAGCCATCAGCAGTTCC
SOCS2	Q164K	CCGCTCTACACGTCAGCACCATCTCTGAAGCATCTCTGTAGGCTCACCATTAAC	GTTAAGTGTGAGCCTACAGAGATGCTTCAGAGATGGTGCTGACGTGTAGAGCGG
	R186S+Y190D	CATCTGGGGACTGCCTTTACCAACAAGCCTAAAAGATGACTTGGAAG	CTTCCAAGTAATCTTTTAGGCTTGTTGGTAAAGGCAGTCCCCAGATG

6.1.4 Transformation

50 μl of chemically *E. coli* competent cells [DH5α strain was used for cloning and BL21(DE3) strain was used for protein expression] were thawed on ice. To these, 100-150 ng of plasmid DNA (1-3 μl) were added and incubated for 30 min on ice. Then, the cells were heat-shocked for 45 sec in a water bath at 42 °C and immediately cooled on ice for 5 min. The cells were diluted by addition of 150 μl of lysogeny broth (LB) media and then incubated in a 37 °C water bath for 60 min (this step was skipped when the antibiotic resistance was ampicillin). Next, the cell suspension was plated on LB/agar plates containing the appropriate antibiotic and incubated at 37 °C overnight.

6.1.5 BL21 pGro7 competent cells generation

BL21(DE3) cells were transformed with the pGro7 plasmid (Chaperone plasmid set, TaKaRa) according to the protocol above. Competent cells were

prepared from a single colony of transformant following the protocol from Sambrook and Russel 2006.[168] The competent cells containing the pGro7 plasmid were aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

6.1.6 Preparation of DH10Bac competent cells

DH10Bac cells from Dr. Axel Knebel were streaked on a LB agar plate supplemented with kanamycin and tetracycline. The plate was incubated overnight at 37 °C. A single colony of DH10Bac was used to inoculate 5 ml of LB supplemented with kanamycin and tetracycline. The starter culture was incubated overnight at 37 °C and 180 rpm. A larger culture of 300 ml of LB with antibiotics was inoculated with 3 ml of the starter culture and incubated at 37 °C, 180 rpm, until reaching an OD_{600} ~0.380. At this point the cells were aliquoted in 50 ml Falcon tubes pre-chilled in ice and incubated in ice for 10 min. The cells were pelleted by centrifugation at 180 x g for 15 min and the pellet was then re-suspended in 10 ml of ice-cold 60 mM $CaCl_2$, 15% (v/v) glycerol, 10 mM PIPES pH 7.0 (adjusted with KOH). This centrifugation and re-suspension procedure was repeated and the cells were incubated in ice for 30 min. After this time, the cells were again pelleted and then re-suspended in 2 ml of $CaCl_2$ solution. The DH10Bac competent cells were aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

6.1.7 Bacmid generation

30 ng (3 μl) of Rbx1–Cul2 plasmid were transformed into DH10Bac competent cells following the protocol described in section 6.1.4. The recovery time was extended to 4 h and then the cells were plated onto LB agar containing tetracycline, kanamycin, X-gal and IPTG. The bacmid propagates in

E. coli DH10Bac as a large plasmid that confers resistance to kanamycin and can complement a lacZ deletion present on the chromosome to form colonies that are blue (Lac+) in the presence of a chromogenic substrate such as X-gal and the inducer, IPTG. The resistance to tetracycline is conferred by the helper plasmid present in the DH10Bac cells. The plates were incubated for 48 h (at least) at 37 °C in order to discriminate between blue and white colonies. 5 white colonies and one blue colony were streaked onto a new LB agar plate containing the adequate supplements to discard false positives. One of the white colonies from the second plate was used to inoculate 300 ml of LB supplemented with tetracycline, gentamycin and kanamycin. The culture was incubated at 37 °C under shaking (130 rpm) for 24 h. Then the cells were pelleted and submitted to the DNA sequencing facility for isolation of the bacmid through maxiprep. The bacmids were stored at 4 °C until transfection and then were aliquoted, flash frozen in liquid nitrogen and stored at -80 °C.

6.1.8 Insect cells tissue culture

Spodoptera fugiperda 21 (*Sf21*) cells were cultured in suspension in Insect-XPRESS media (Lonza) supplemented with amphotericin B, penicillin and streptomycin (Antimycotic-antibiotic 100x solution, Gibco). The cultures were incubated at 26.5 °C with 130 rpm shaking and in the dark. Every 48-72 h the cells were split to a density of 1.5 x 10^6 cells/ml in fresh media.

6.1.9 Bacmid transfection into *Sf21* cells

The transfections were performed in adherent cell culture. To do that, 1.5 x 10^6 cells were seeded in a T25 flask 30 min prior to the transfection. The transfection mixture was prepared in a 24-well plate, where 500 μl of media

without antibiotic, 15 μl of Cellfectin (ThermoScientific) and 1-3 μg of bacmid were added to a well. The mixture was incubated at room temperature, in the dark, for 30-60 min. After this period, 1.5 ml of media were added to the well and the solution was mixed by gently pipetting up and down. Then the solution was transferred to the T25 flask with the cells (the media in the flask must be removed first) and media was added to make up 5 ml. The flask with the cells was incubated overnight at 26.5 °C. The morning after, the transfection mixture was aspirated and 7 ml of fresh media supplemented with antimycotic-antibiotic solution were added to the flask. The flask was incubated at 26.5 °C for 7 days. Finally, the cells were harvested by scrapping the flask, the solution was transferred to a falcon tube and the cells were spun down (180 x g, 4 min). The supernatant (P0 virus) was stored in a new, sterile, falcon tube at 4 °C in the dark. The efficiency of the transfection was assessed by analysing the presence of Cul2 in the cell pellet by immunoblot with anti-Cul2 (Abcam, Ab166917) at 1:5000 dilution.

6.1.10 Baculovirus amplification

The baculovirus was amplified in suspension cell culture. 150 ml of cells at a density of 1.5 x 10^6 cells/ml were infected with 1.5 ml of P0 virus. The culture was incubated for 5 days at 26.5 °C, 130 rpm shaking in the dark. After this period the cells were harvested by centrifugation (180 x g, 10 min) and the supernatant (P1) was stored at 4 °C in the dark.

6.1.11 Protein expression in bacteria

VBC and SBC ternary complexes and variant proteins were purified as described previously.[52,116] BL21(DE3) *E. coli* cells were co-transformed with the

plasmid for expression of pVHL/SOCS2 and the bi-cistronic pDUET plasmid for expression of EloBC. A single colony of transformant was used to inoculate LB media for bacterial culture. Protein expression was induced with 0.3 mM IPTG (when OD_{600} reached 0.8) at 24 °C for 18 h. Co-expression of these proteins resulted in the formation of the respective trimeric complex (VBC/SBC) that was then purified by two steps of affinity chromatography, followed by ion-exchange chromatography and finally by size-exclusion chromatography. The His-tag was cleaved between the two affinity chromatography steps with TEV protease. Following this protocol the yield of protein was about 15-20 mg per litre of culture.

$Cul5_{NTD}$ (residues 1−386) was also expressed in BL21(DE3) *E. coli* cells.[115] After transformation and inoculation of LB media for bacterial growth, protein expression was induced with 0.5 mM IPTG at 18 °C for 18 h. His-tagged $Cul5_{NTD}$ was purified by affinity chromatography and by size-exclusion chromatography with a yield of ~40 mg of protein per litre of culture.

For the $Cul2_{NTD}$ constructs, a single colony from BL21(DE3) *E. coli* cells transformed with the respective plasmid was inoculated in 100 ml of LB media with ampicillin and incubated overnight at 37 °C and 180 rpm shaking (starting culture). The following day, the starting culture was used to inoculate Terrific Broth (TB) media supplemented with ampicillin, tetracycline and arabinose for large-scale expression to an OD_{600} of 0.05. Cell cultures were incubated at 37 °C and 180 rpm shaking until reaching an OD_{600} of 0.6 and then protein expression was induced with the appropriate IPTG concentration (see Chapter 3) and the temperature was lowered to 16 °C. The cultures were incubated in these conditions for 16-20 h. Then, cells were harvested by centrifugation (30 min at 5,000 x g), the cell pellet was re-suspended in lysis

buffer (supplemented with DNAse, $MgCl_2$ and protease inhibitors) and the cells were lysed by French press. The cell debris was removed by centrifugation (20 min at 50,000 x g) and the supernatant was applied into an MBP-trap column. The protein of interest was eluted from the column with a gradient of maltose. Next, the sample was applied in a gel filtration column (analytical, Superdex Increase 10/300, or preparative HiLoad 16/600 Superdex 200).

The protein's identities were confirmed by electrospray mass spectrometry analysis (Figure 6.2 – 6.8).

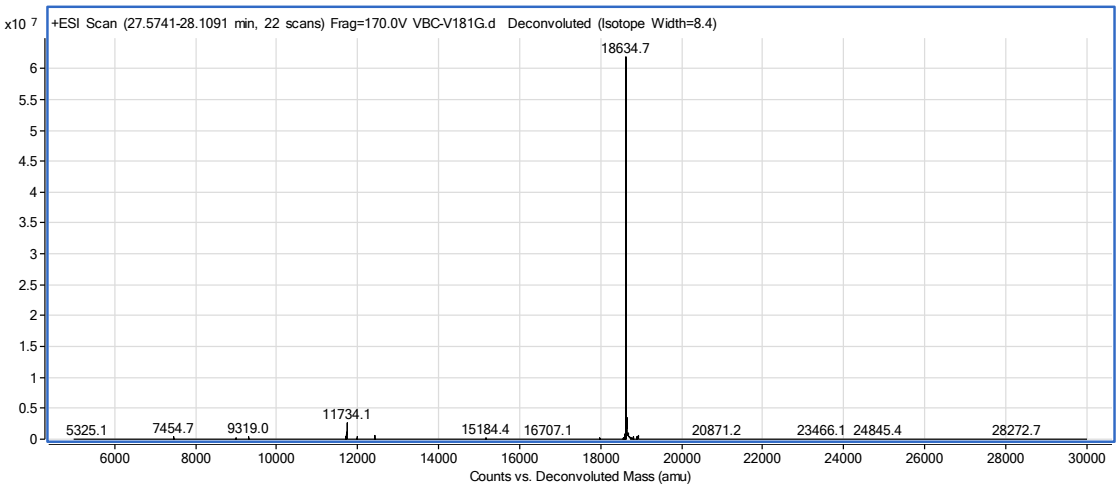


Figure 6.2 – Deconvoluted electrospray MS spectrum of pVHL V181G from purified VBC.

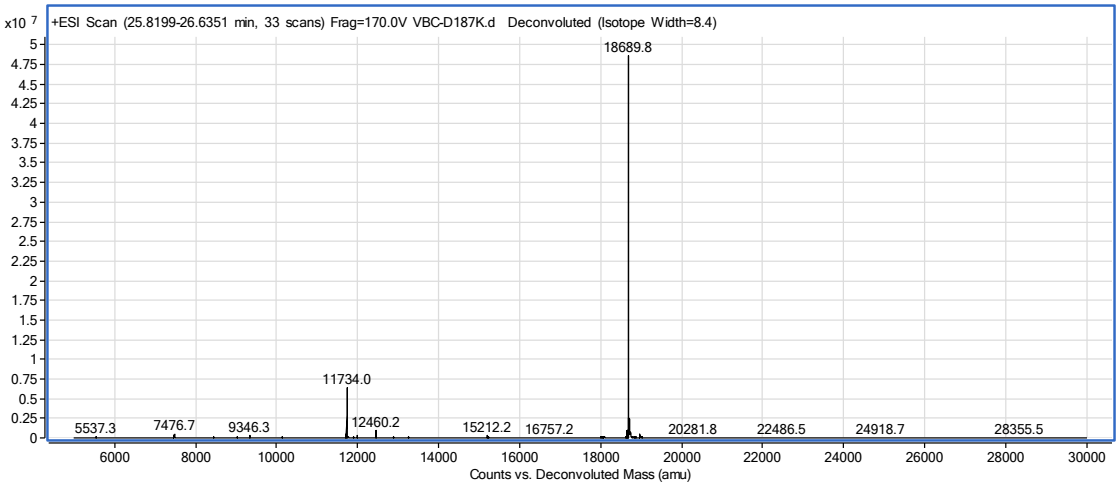


Figure 6.3 – Deconvoluted electrospray MS spectrum of pVHL D187K from purified VBC.

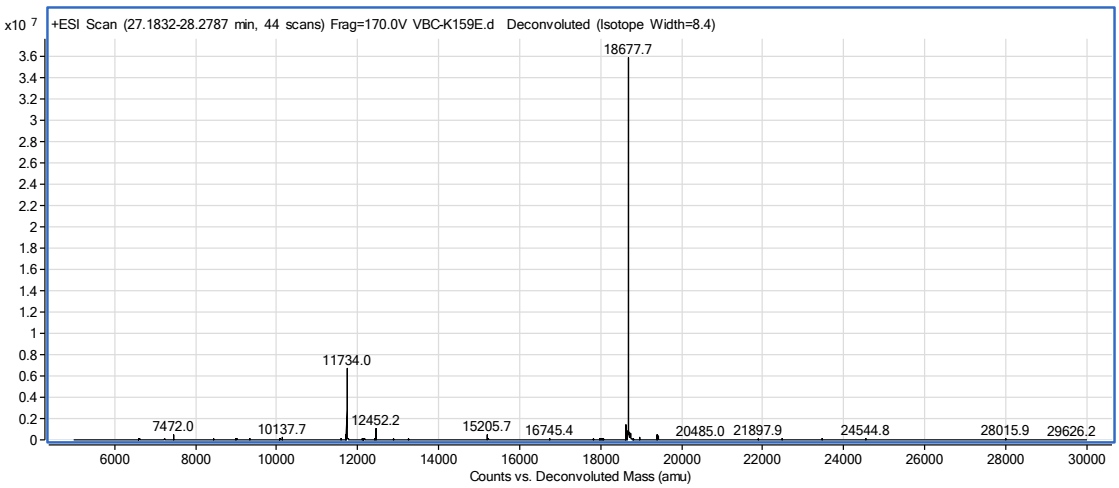


Figure 6.4 – Deconvoluted electrospray MS spectrum of pVHL K159E from purified VBC.

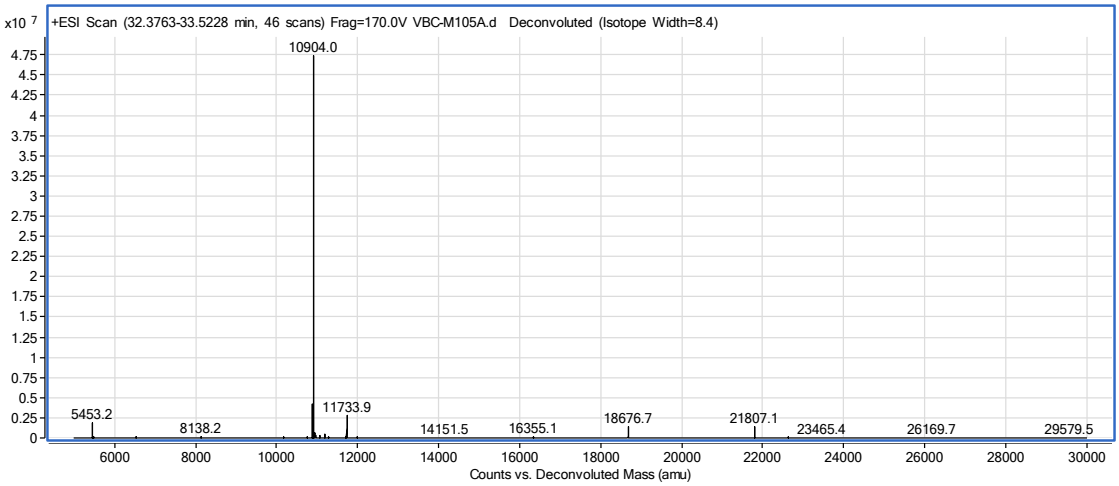


Figure 6.5 – Deconvoluted electrospray MS spectrum of EloC M105A from purified VBC.

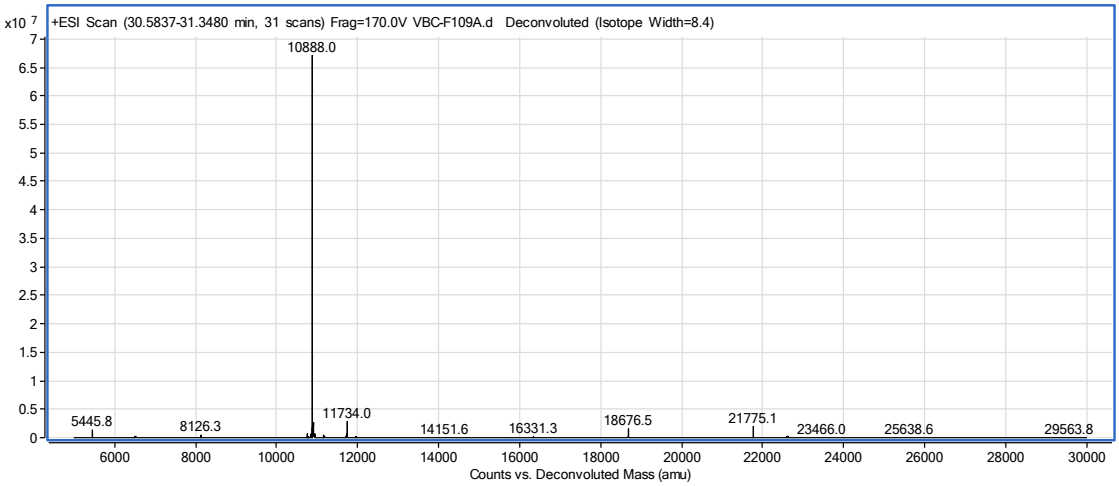


Figure 6.6 – Deconvoluted electrospray MS spectrum of EloC F109A from purified VBC.

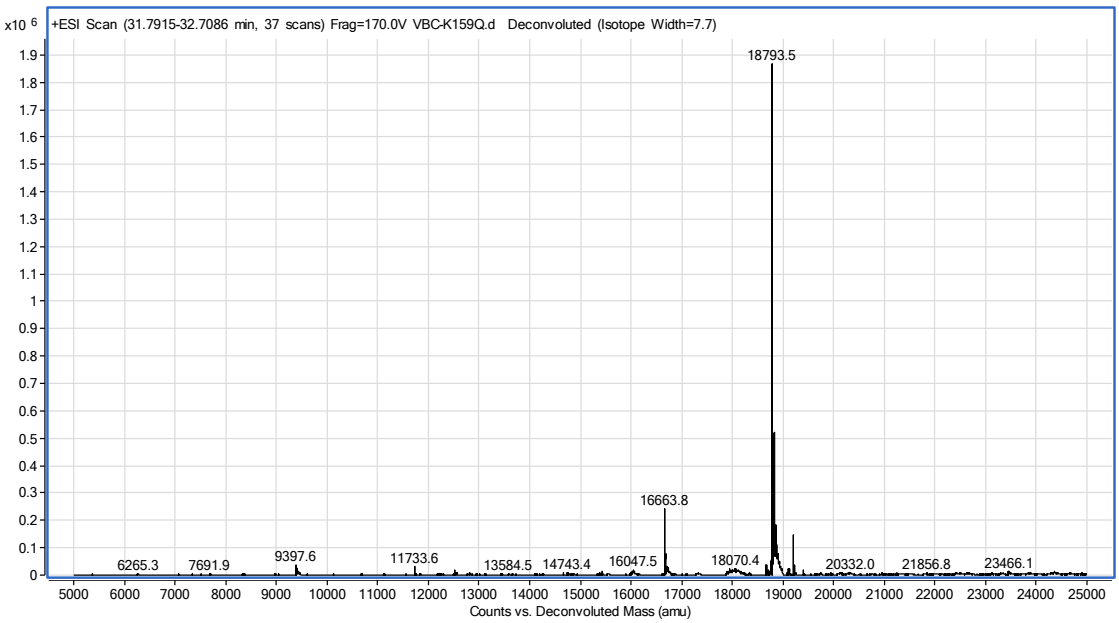


Figure 6.7 – Deconvoluted electrospray MS spectrum of pVHL K159Q D187Y S183R from purified VBC.

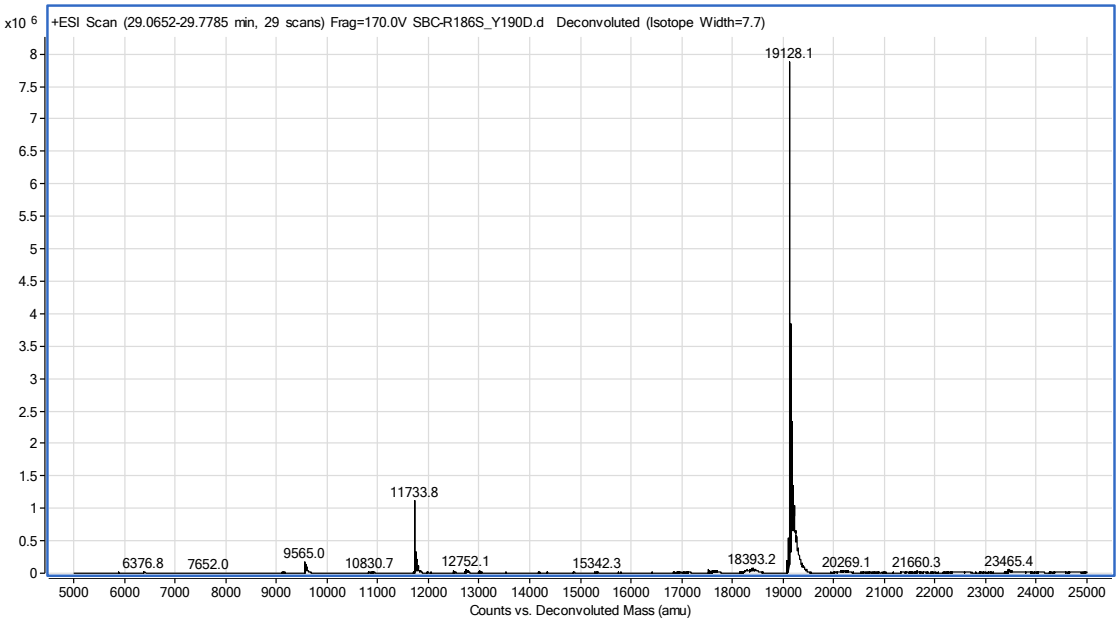


Figure 6.8 – Deconvoluted electrospray MS spectrum of SOCS2 Q164K Y190D R186S from purified SBC.

6.1.12 Protein expression in insect cells

Sf21 cells at a density of 1.5×10^6 cells/ml were infected with the P1 virus in a 1:100 ratio and incubated at 26.5°C, 130 rpm, in the dark for 72 h. The cells were harvested by centrifugation, the pellet was re-suspended in lysis buffer

containing 50 mM HEPES pH 8.0, 250 mM NaCl, 2 mM TCEP and 0.2% (v/v) Triton-X and supplemented with DNAse, $MgCl_2$ and protease inhibitors (leupeptin and Pefabloc). The cells were lysed by French press, the lysate was clarified by centrifugation and the supernatant was mixed with ampicillin-modified sepharose resin. As mentioned earlier, Cul2 contains a N-terminal Dac tag, which binds ampicillin. Therefore, the affinity purification is performed using ampicillin-modified resin (see section 6.1.18). After 1 h incubation at room temperature, the resin was washed three times with 20 mM HEPES pH 8.0, 100 mM NaCl, 5% (v/v) glycerol, 2 mM TCEP. At the last wash step the resin was suspended in the same buffer and incubated with TEV enzyme for 2.5 h at room temperature. The cleaved Rbx1–Cul2 was recovered through filtration and the filtrate, after concentration, was loaded on a Superdex 200 gel filtration column (GE Healthcare) for further purification.

The protein's identities were confirmed by electrospray mass spectrometry analysis (Figure 6.9 – 6.18).

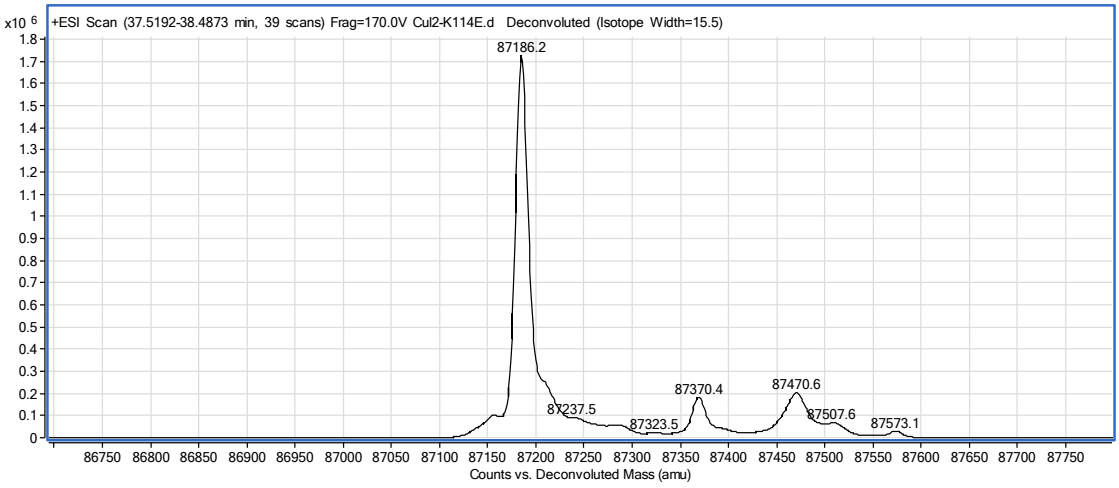


Figure 6.9 – Deconvoluted electrospray MS spectrum of purified Cul2 K114E from purified Rbx1–Cul2.

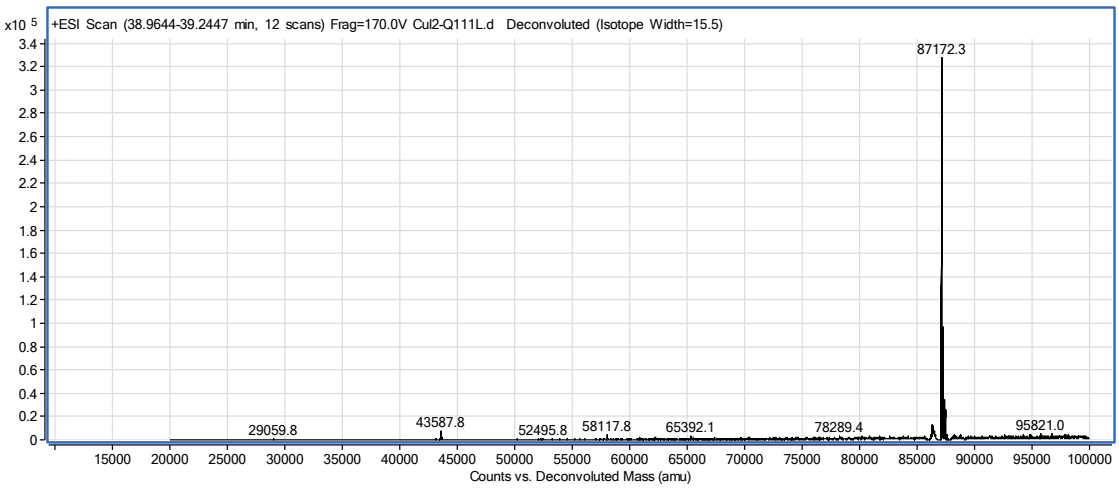


Figure 6.10 – Deconvoluted electrospray MS spectrum of Cul2 Q111L from purified Rbx1–Cul2.

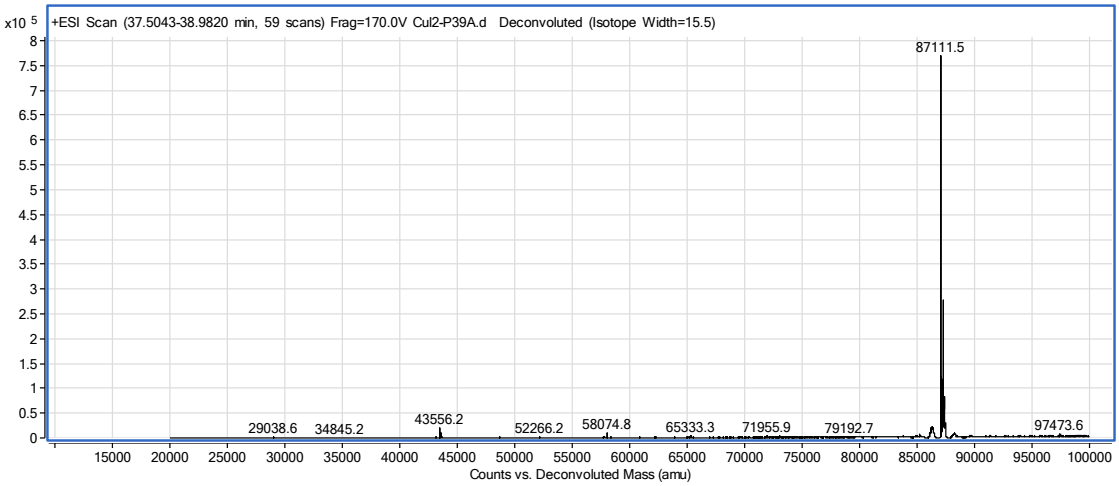


Figure 6.11 – Deconvoluted electrospray MS spectrum of Cul2 F39A from purified Rbx1–Cul2.

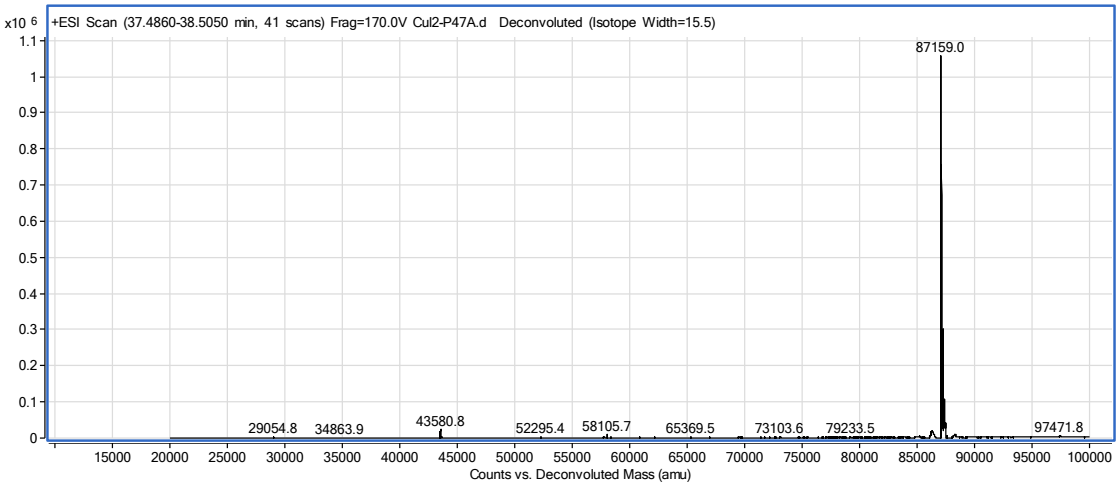


Figure 6.12 – Deconvoluted electrospray MS spectrum of Cul2 V47A from purified Rbx1–Cul2.

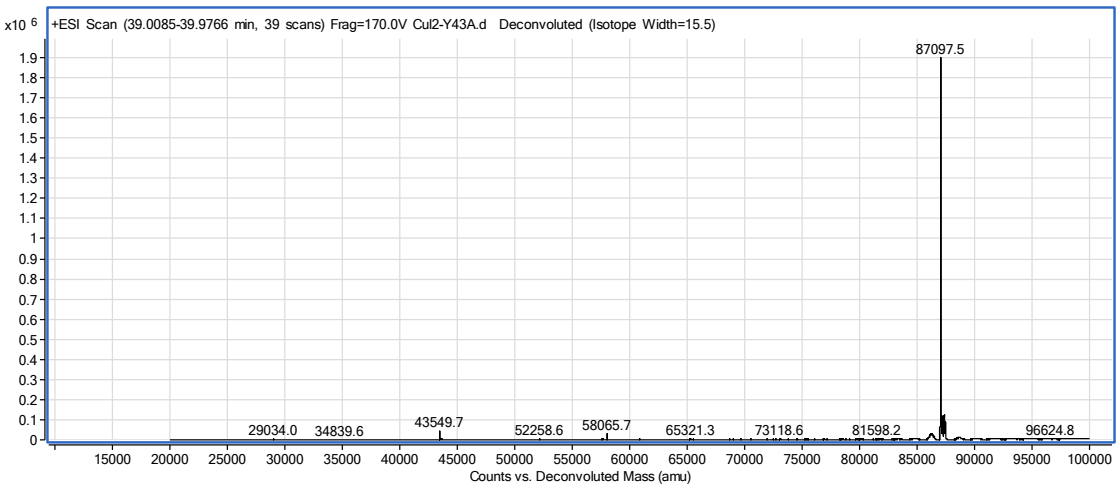


Figure 6.13 – Deconvoluted electrospray MS spectrum of Cul2 Y43A from purified Rbx1–Cul2.

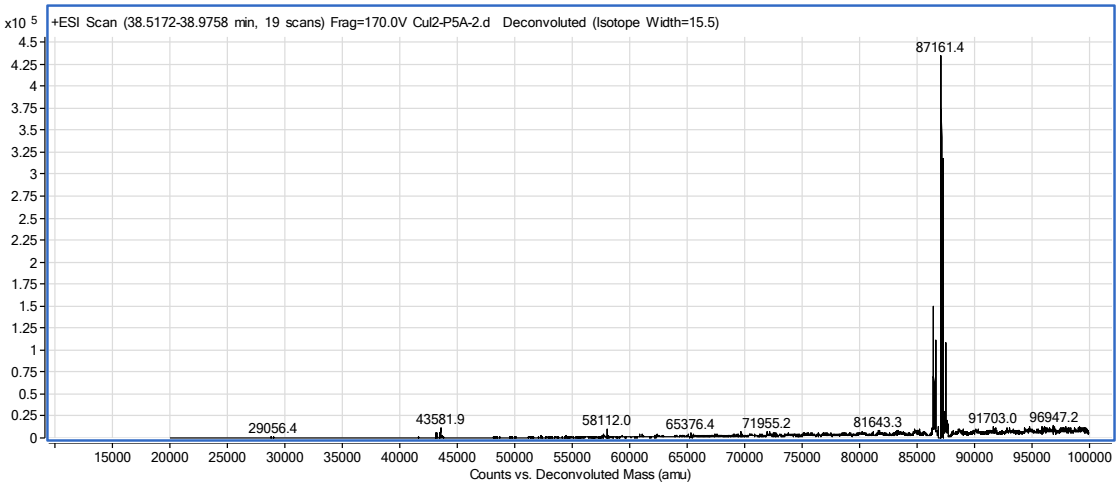


Figure 6.14 – Deconvoluted electrospray MS spectrum of Cul2 P5A from purified Rbx1–Cul2.

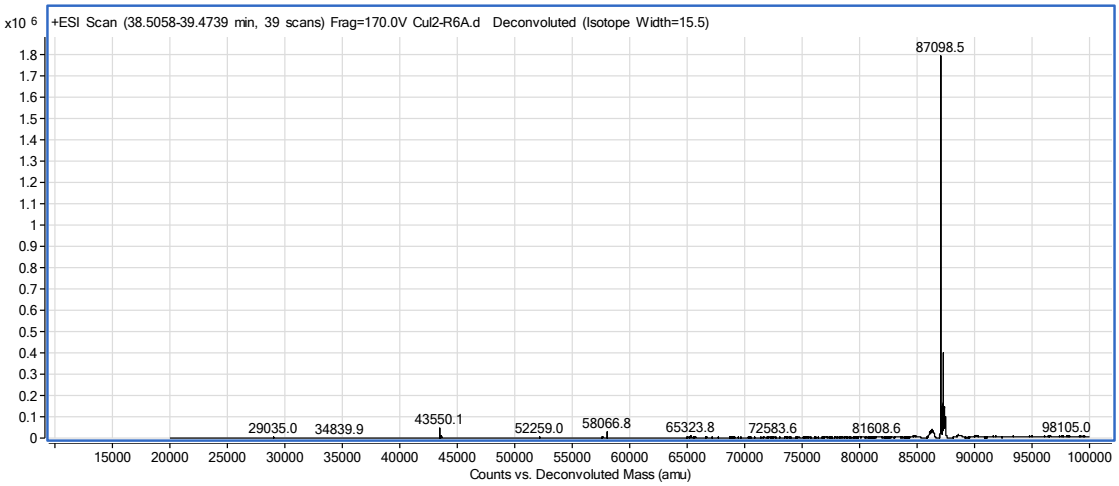


Figure 6.15 – Deconvoluted electrospray MS spectrum of Cul2 R6A from purified Rbx1–Cul2.

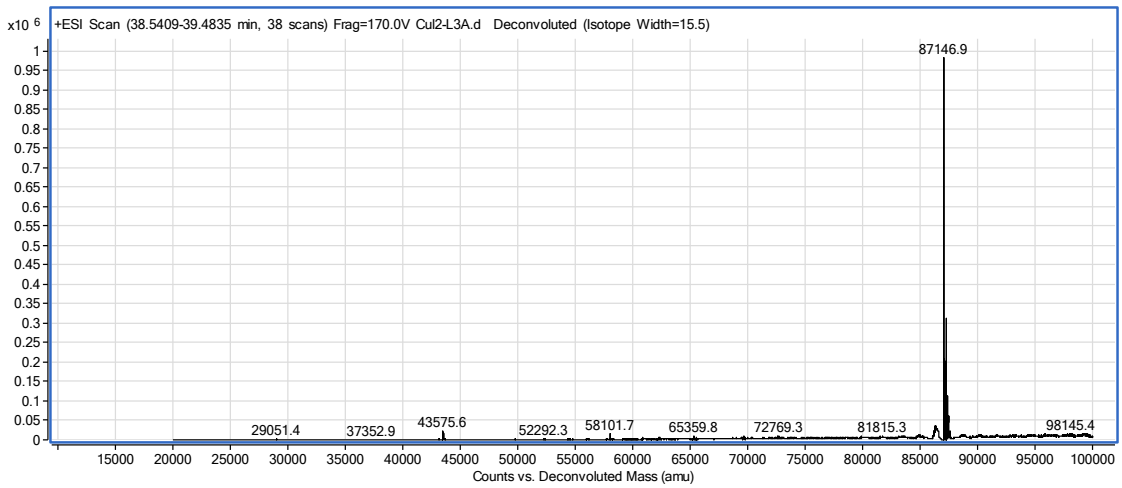


Figure 6.16 – Deconvoluted electrospray MS spectrum of Cul2 L3A from purified Rbx1–Cul2.

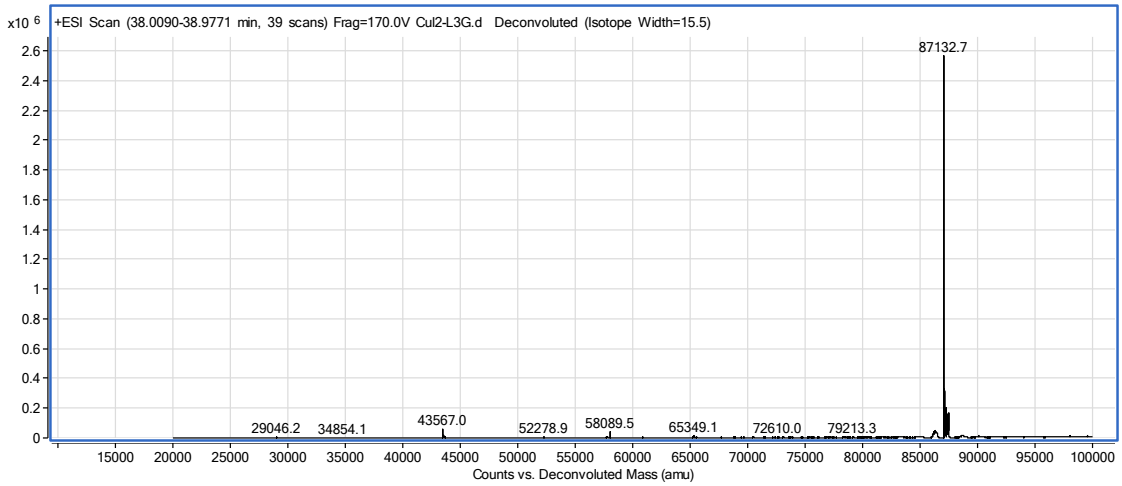


Figure 6.17 – Deconvoluted electrospray MS spectrum of Cul2 L3G from purified Rbx1–Cul2.

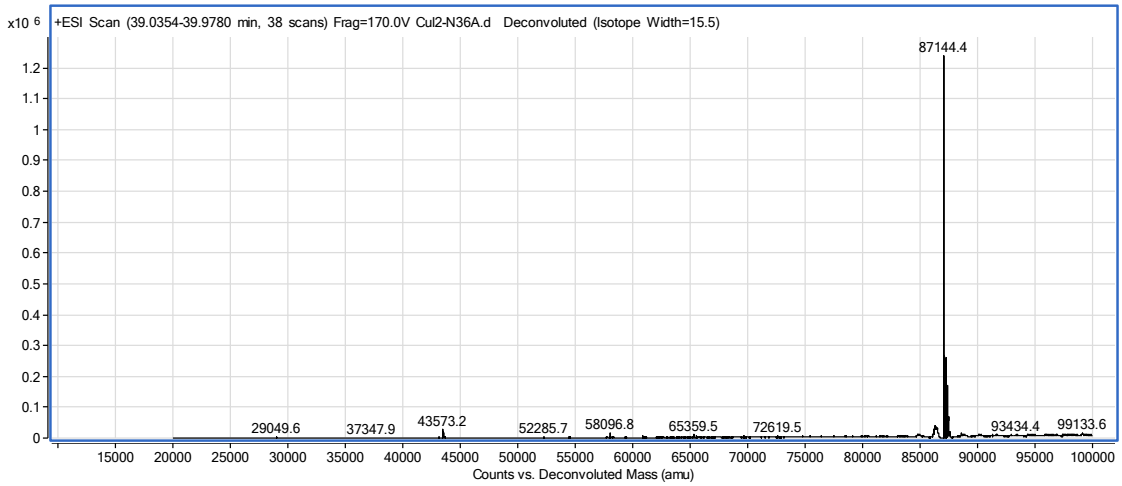


Figure 6.18 – Deconvoluted electrospray MS spectrum of Cul2 N36A from purified Rbx1–Cul2.

6.1.13 Isotopically-labelled VBC expression

Protein expression was performed in ^{2}H, ^{15}N enriched media (*E.coli*-OD2, Silantes). The cell cultures were incubated at 37 °C until reaching an OD_{600} of 0.8 (13 h) and then protein expression was induced with 0.6 mM of IPTG and the cultures were incubated for 16 h at 24 °C (OD_{600} ~ 2). The protein was purified according to the protocol previously described in section 6.1.11. Considering the purity of the sample achieved, the gel filtration step was skipped. The protein was obtained with a yield of 4 mg/litre. The success of the isotopic enrichment of the sample was assessed by mass spectrometry (Figure 6.19). The percentage of deuteration was calculated assuming the total incorporation of ^{15}N, and it was determined to be 99.1% for pVHL, 99.0% for EloC and 98.9% for EloB.

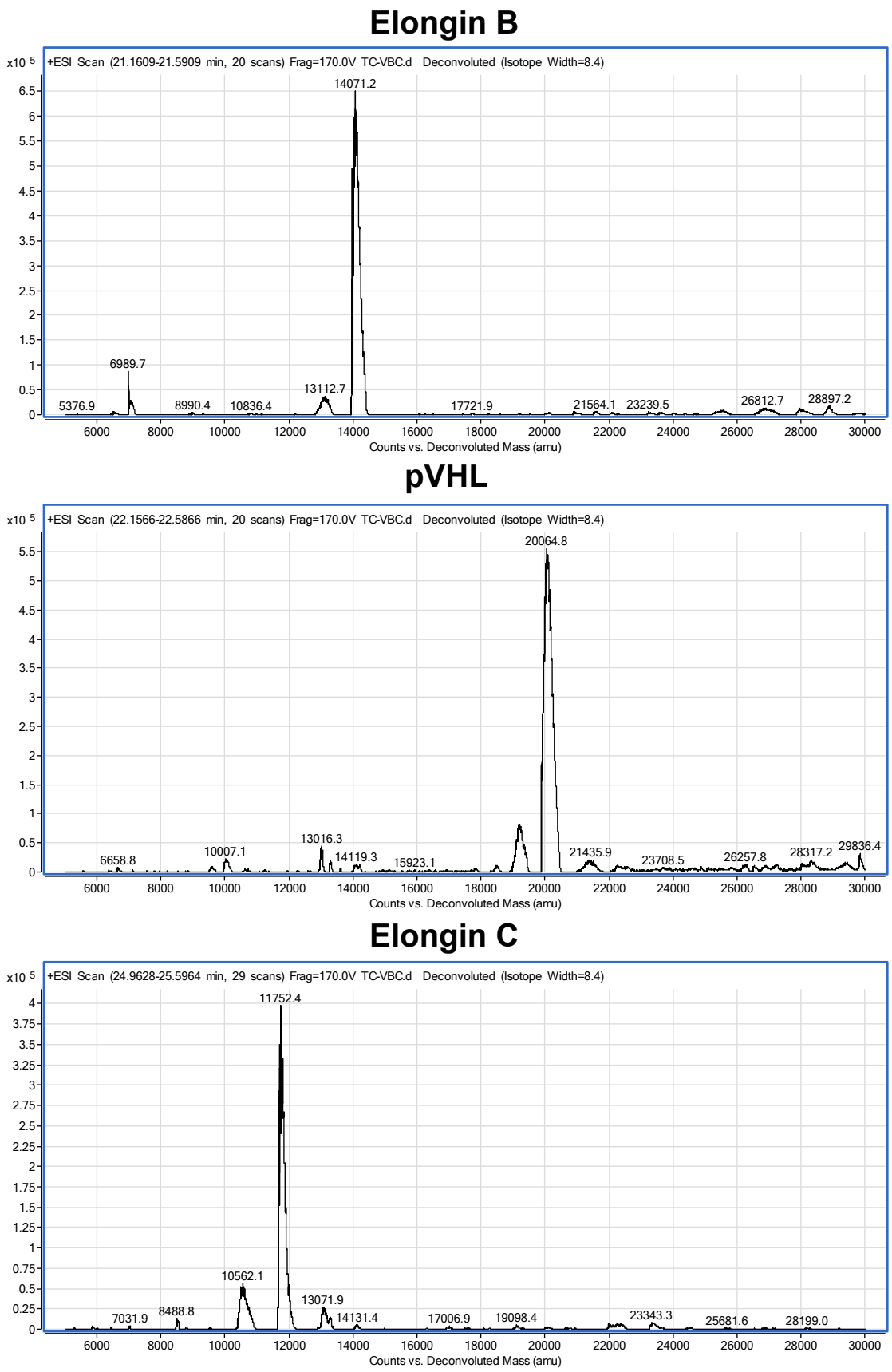


Figure 6.19 – Electrospray MS of purified $^{2}H,^{15}N$-VBC. Spectra for Elongin B, pVHL and Elongin C. The expected mass accounting for fully incorporation of ^{15}N and ^{2}H was 14220.8 Da for EloB, 20264.3 Da for pVHL and 11860.9 Da for EloC.

6.1.14 Biotinylation of proteins

Proteins were biotinylated using the EZ-link NHS-biotin kit (ThermoFisher Scientific) and following the manufacturer guidelines. Briefly, the protein was dialysed into a buffer that did not contain primary amines and then mixed with biotin in equimolar ratio and incubated for 30 min at room temperature. The excess of biotin was removed by desalting column (Zeba Spin desalting column 0.5 ml, ThermoFisher Scientific).

6.1.15 DNA and protein concentrations

The concentration of DNA and protein samples was estimated measuring the absorbance at 260 nm and 280 nm, respectively, using a Nanodrop Microliter UV/Vis spectrophotometer (ThermoScientific). DNA purity was assessed by the A_{260}/A_{280} ratio and in all the samples used $A_{260}/A_{280} > 1.7$. Protein concentration was calculated based on the theoretical molar extinction coefficient predicted using ProtParam.[169]

6.1.16 Immunoblotting

The cells lysates were run on a SDS-PAGE and transferred to a nitrocellulose membrane using i-Blot 2® (Life Technologies). The membrane was blocked with 10% of BSA in Tris-buffered saline (TBS) with 0.1% (v/v) Tween-20 (TBST) for 1 h at room temperature. Then, the membrane was incubated overnight at 4 °C with the primary antibody - anti-Cul2 (Abcam, Ab166917) at 1:5000 dilution in TBST. The membrane was brought to room temperature for 10-30 min and then washed three times with TBST (15 min, 5 min, 5 min). After this step the membrane was incubated with the secondary antibody - anti-rabbit (Abcam, Ab191866) at a dilution of 1:2000 in 5% BSA in

TBST for 1 h at room temperature. The washing steps were repeated and the protein detection was performed using the Pierce ECL chemiluminescence kit (Thermo Scientific).

6.1.17 Neddylation assay

The neddylation assay was performed according to the protocol provided by Dr. Axel Knebel (MRC-PPU).

The assay buffer consisted on 30 mM HEPES pH 7.5. 100 mM NaCl, 1 mM DTT and 100 mM MgCl2, 2 mM ATP (both solutions were prepared as 10x solution). 10 μg of protein (Rbx1–Cul2, $CRL2^{VHL}$ or $CRL2^{VHL:HIF}$) were mixed with 1.2 μg of E1 enzyme, 2 μg of E2 enzyme and 13 μg of NEDD8 in a total volume of 25 μl. The reaction was prepared in PCR tubes and those were incubated at 37 °C for 2 h. The reaction was stopped by adding 4x SDS buffer and heating up the samples at 95 °C for 10 min. The samples were then run on two SDS-PAGE - one for staining with Coumassie and other for immunoblotting with anti-Cul2.

6.1.18 Preparation of ampicillin-modified sepharose resin

25 ml of resin (NHS-activated sepharose 4 Fast-flow, GE Healthcare) were poured on a filtering device (e.g. stericup, column filter). The resin was washed abundantly with 1 mM HCl (> 500 ml), and then transferred to a 50 ml Falcon tube. The excess of HCl solution was removed by spinning down the resin and pouring out the liquid. Next, the resin was re-suspended in 25 ml of 0.4 M ampicillin in 0.1 M $NaHCO_3$ pH 7.5 and incubated in a roller at room temperature for 1-2 h. After this, the excess of ampicillin was removed by washing the resin thoroughly with any buffer at pH < 7.0 (in the filter device).

Finally, the resin was washed several times with 30% (v/v) ethanol and transferred to a 50 ml Falcon tube. The excess of liquid was removed as before and the resin was re-suspended in an equal volume of 30% (v/v) ethanol solution (50% slurry). The ampicillin-modified resin was stored at 4 °C (stable up to 6 months).

6.2 Crystallography

6.2.1 $CRL2^{VHL}$ complex crystallisation

VBC was incubated with a HIF-1α 19-mer peptide (residues 559-557) and the resulting complex (VBCH) was purified on a Superdex 75 gel filtration column (GE Healthcare) as previously described.[52] Next, VBCH and Rbx1–Cul2 were mixed in equimolar ratio and incubated for 30 min at room temperature. The $CRL2^{VHL}$ complex was purified in an analytical Superdex 200 Increase gel filtration column (GE Healthcare) equilibrated with 50 mM Tris-HCl, 150 mM NaCl, 5 mM dithiothreitol (DTT). The pentameric complex was concentrated to 4.2 mg/ml. Equal volumes of $CRL2^{VHL:HIF}$ complex and liquor solution were mixed in the hanging-drop vapour diffusion method at 20 °C. The liquor solution contained 0.1 M Tris pH 7.6, 0.15 M ammonium sulphate, 15% polyethyleneglycol 4000 and 3% 1,4-dioxane or 4% acetonitrile as additive. After equilibration, the drop was streaked with seeds of disrupted $CRL2^{VHL:HIF}$ crystals. Crystals would generally appear within 48 h. Crystals were cryoprotected with 20% ethyleneglycol or 20% glycerol and screened for diffraction patterns using an in-house Rigaku M007HF X-ray generator and Saturn 944HG+ CCD detector.

6.2.1.1 Data collection and structure solving

X-ray data were collected at 100 K and a wavelength of 0.9282 Å at Diamond Light Source beamline I04-1. Indexing and integration of reflections was performed using DIALS, and scaling and merging with AIMLESS in CCP4i.[170,171] The isomorphous datasets were refined using REFMAC5[98] and COOT.[172] The MOLPROBITY[173] server was used to validate the geometry and steric clashes in the structures. The Ramachandran analysis indicated 90.7% residues in preferred regions, 9.1% residues in allowed regions and 0.2% outliers. The data collection and refinement statistics are presented in Table 2.1 in Chapter 2.

6.2.2 VBC crystallisation and soaking

The VBC complex was crystallised according to the protocol described previously.[58] The crystallisation solutions contained 0.1 mM sodium cacodylate, pH 6.2–6.5, 16–18% polyethylene glycol 3350, 0.2 M magnesium acetate and 10 mM DTT. VBC and crystallisation solution were mixed in equal volumes and crystals were grown using the hanging-drop vapour diffusion method at 18 °C. To slow the vapour diffusion rate, a thin layer of Al's Oil (Hampton Research) was applied on top of crystallisation solutions. The drops were streak seeded with crushed VBC crystals grown in the same conditions originating well-diffracting crystals.

Peptides were soaked overnight in 1–100 mM solutions of ligand in 1-10% dimethylsulfoxide, 99–90% liquor solution. Crystals were cryo-protected in 20% glycerol and flash frozen in liquid nitrogen. Crystals were screened using an in-house Rigaku M007HF X-ray generator and Saturn 944HG+ CCD detector.

6.3 Phage display

6.3.1 Library amplification

The phage display libraries VL1, VL2 and VL3 were a gift from Dr. Christian Heinis (EPFL Lausanne). New stocks of the libraries were prepared by re-suspending the content of each tube in 10 ml of 2xYT media supplemented with chloramphenicol and plating the cells in large 2xYT/agar plates. The cells were incubated overnight at 37 °C. The following day the cells were scrapped from the surface of the plates with 5 ml of 2xYT media. 15% glycerol was added and each library was dispensed into 0.5 ml aliquots, flash frozen in liquid nitrogen and stored at -80 °C.

Before each panning experiment, the respective library was amplified. To do this, 2 x 500 ml of 2xYT media supplemented with chloramphenicol (in 2 L flasks) were inoculated with 2 x 0.5 ml of the corresponding library and incubated overnight at 30 °C with 250 rpm shaking. Then, the cells were centrifuged at 5,000 x g for 30 min at 4 °C and the phage-containing supernatant was transferred to a new tube containing ¼ (v/v) solution of ice-cold 20% PEG-6000 and 2.5 M NaCl. After inverting the tube several times, the phage solution was kept in ice for 30 min to precipitate the phages. After this time, the solution was centrifuged for another 30 min (5,000 x g, 4 °C), the supernatant was discarded and the excess of liquid was removed with a filter paper. The precipitated phages were re-suspended in 20 ml reaction buffer (Table 6.4) and stored in a 50 ml Falcon tube at 4 °C.

6.3.2 Peptide cyclisation

Initially, the cysteine residues were reduced by adding 1 ml of 20 mM tris-(2-carboxyethyl)phosphine (TCEP) to 19 ml of phage solution and incubated

for 1 h at 42 °C. After this time, the solution was cooled down by incubation in ice for about 5 min and then transferred to a 100,000 Da centricon. The volume of the solution was concentrated to 1 ml and 10 ml of ice-cold reaction buffer were used to re-suspend the phage. This procedure was repeated a couple of times and finally the solution was transferred to a 50 ml tube, where the volume was adjusted to 32 ml with reaction buffer.

To cyclise the peptides, 8 ml of 50 μM TBMB in acetonitrile were added to the 32 ml of phage solution and the tube was incubated at 30 °C for 1 hour. Then, the phages were precipitated with 10 ml of ice-cold PEG/NaCl solution and incubated in ice for 15 min. The solution was centrifuged for 30 min at 5,000 x g, 4 °C, the supernatant was discarded and the precipitate was re-suspended in 3 ml of binding buffer (Table 6.4). The phages were stored at 4 °C until used for phage selection (stable for 3-5 days).

6.3.3 Biopanning

VBC/VBCH (4-5 μg) were immobilised in tosyl-activated magnetic beads (Dynabeads M-280 tosyl-activated, Invitrogen) following the protocol of the manufacturer. Unbound protein was washed with binding buffer (Table 6.4). The phages and the protein attached to the beads were incubated with blocking buffer (Table 6.4) for 30 min on the slow rotating wheel at room temperature, the phages were added to the protein-beads and incubated again for another 30 min. The beads were washed several times with washing buffer (Table 6.4) and twice with binding buffer before elution with 100 μl of elution buffer (Table 6.4) for 5 min and then neutralised in 50 μl of 1 M Tris-HCl pH 8.0. The eluted phages were used to infect 25 ml of *E. coli* cells (ER2738 strain, this K-12 strain expresses the F pilus and is suited for phage propagation and

production; it is an amber suppressor strain, translating the stop codon UAG as glutamine) at OD_{600} between 0.3-0.5 for 90 min at 37 °C. Finally the cells were plated in large 2xYT/chloramphenicol plates. The selection rounds were repeated 3-4 times and in the last round 30 colonies of *E. coli* infected with the phages were submitted to DNA sequencing. A sample from each round of panning was aliquoted and stored in 20% glycerol at -80 °C.

Table 6.4 – Phage display buffers.

Reaction buffer	20 mM NH_4HCO_3 pH 8.0, 5 mM EDTA
Binding buffer	10 mM Tris-Cl pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$ and 1 mM $CaCl_2$
Blocking buffer	binding buffer containing 0.3% (v/v) Tween-20 and 3% (w/v) BSA
Washing buffer	binding buffer containing 0.1% (v/v) Tween-20
Elution buffer	50 mM glycine pH 2.2
Neutralisation buffer	1 M Tris-Cl pH 8.0

6.4 Peptide synthesis and characterisation

Peptides were synthesised in an INTAVIS ResPepSL automated peptide synthesiser by standard solid-phase synthesis using Fmoc-protected amino acids (5 eq., Novabiochem) and Rink amide AM resin (1 eq., 0.4-0.8 mmole/g, Novabiochem). The peptides were synthesised in a 96-well plate set up (bicyclic peptides) or in a 24-column set up. Each cycle started with Fmoc deprotection with 20% piperidine in DMF, followed by two coupling reactions, the first one using 1 eq. of 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and the second one using 1 eq. of 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU) as coupling reagents. Unreacted amino acids were capped with 5% acetic anhydride and N-methylmorpholine (NMM) in DCM

and washed off with DMF. At the last cycle, a final deprotection step was performed and the resin was washed with DMF and dichloromethane (DCM).

Cleavage from the resin and deprotection of the amino acid side chains was performed with a solution of 92.5:5:2.5 (v/v) TFA:triisopropylsilane:H_2O or 90:2.5:2.5:2.5:2.5 (v/v) TFA:H_2O:phenol:thioanisole:1,2-ethanedithiol (bicyclic peptides) for 3 h at room temperature. Then, the cleaved peptides were filtered through a single-fritted column and precipitated with cold diethyl ether. The thin white precipitate was filtered and washed several times with diethyl ether and finally dissolved in water or acetonitrile:water mixture and lyophilised in a centrifugal evaporator (Genevac EZ-2 series, SP Scientific). The peptides were then purified by high-pressure liquid chromatography (HPLC) in the most adequate conditions (described in each section).

6.4.1 Cullin-2 peptides, 15-mer peptide and clamp peptides

The peptides were dissolved in water and purified on preparative HPLC System with a Waters X-Bridge C18 column (100 mm x 19 mm; 5 µm particle size) and a gradient from 5% to 95% of acetonitrile in water over 20 min, under a flow of 25 ml/min, with 0.1% ammonia in the aqueous phase. The fractions collected were lyophilised in the centrifugal evaporator (Genevac EZ-2 series, SP Scientific) and the identity and purity of the peptides was confirmed by analytical HPLC-MS on an Agilent Technologies 1200 series HPLC connected to an Agilent Technologies 6130 quadrupole LC/MS linked to an Agilent diode array detector (Figure 6.20 – 6.47). The column used was an Agilent ZORBAX StableBond 80Å C18, 4.6 x 250 mm, 5 µm column and the compounds were eluted with a gradient of 5−95% acetonitrile/water + 0.1% ammonia. The clamp peptides were purified and analysed in acidic conditions (0.1% TFA).

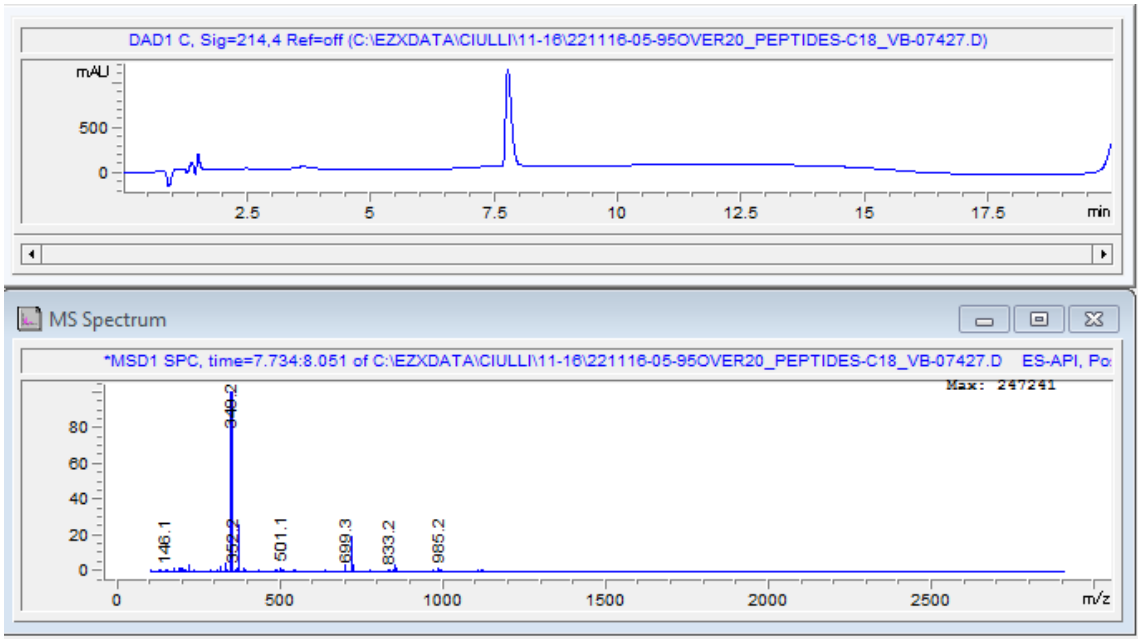


Figure 6.20 – Analytical LC-MS trace of peptide MSL (UV trace recorded at 214 nm).

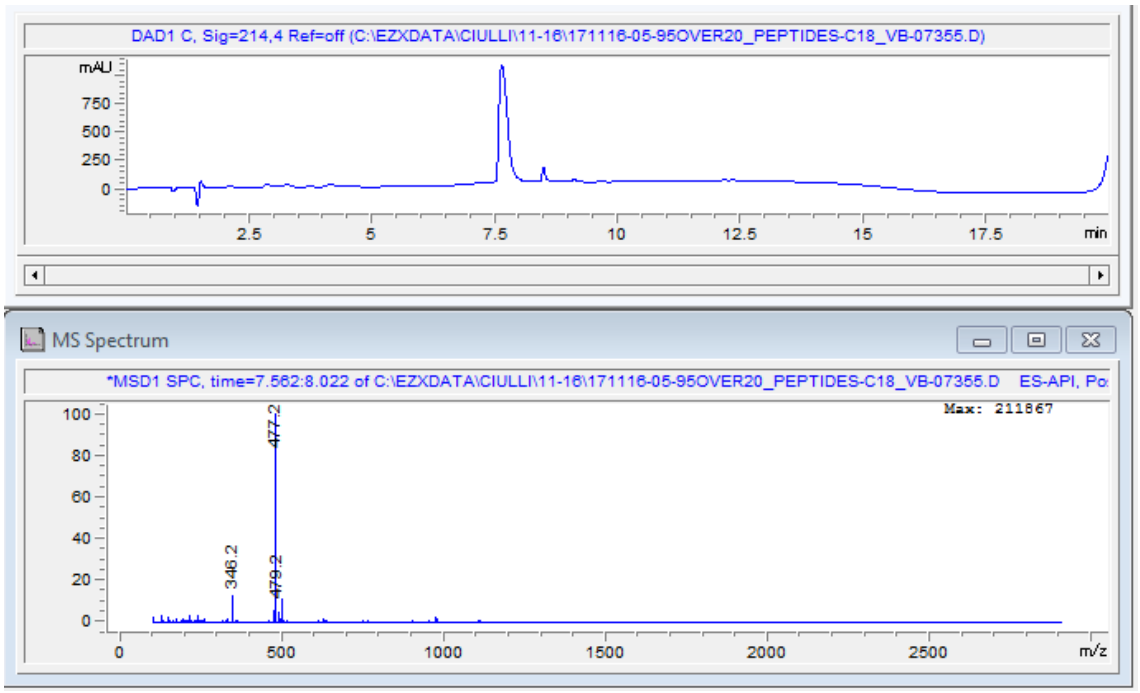


Figure 6.21 – Analytical LC-MS trace of peptide MSLK (UV trace recorded at 214 nm).

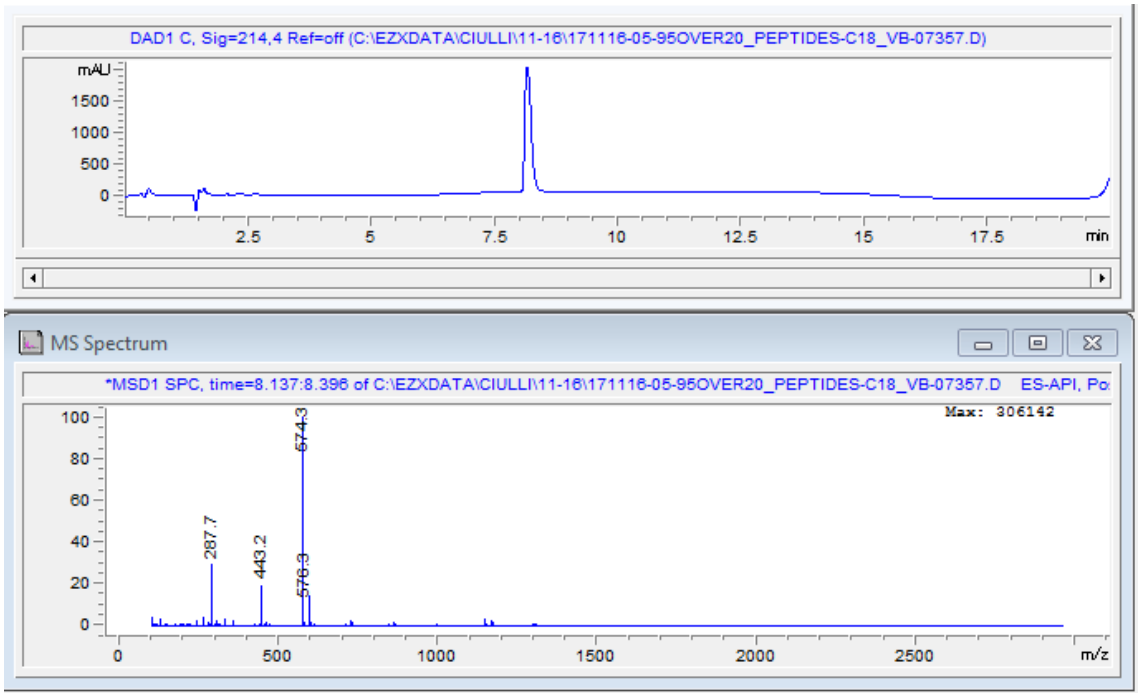


Figure 6.22 – Analytical LC-MS trace of peptide MSLKP (UV trace recorded at 214 nm).

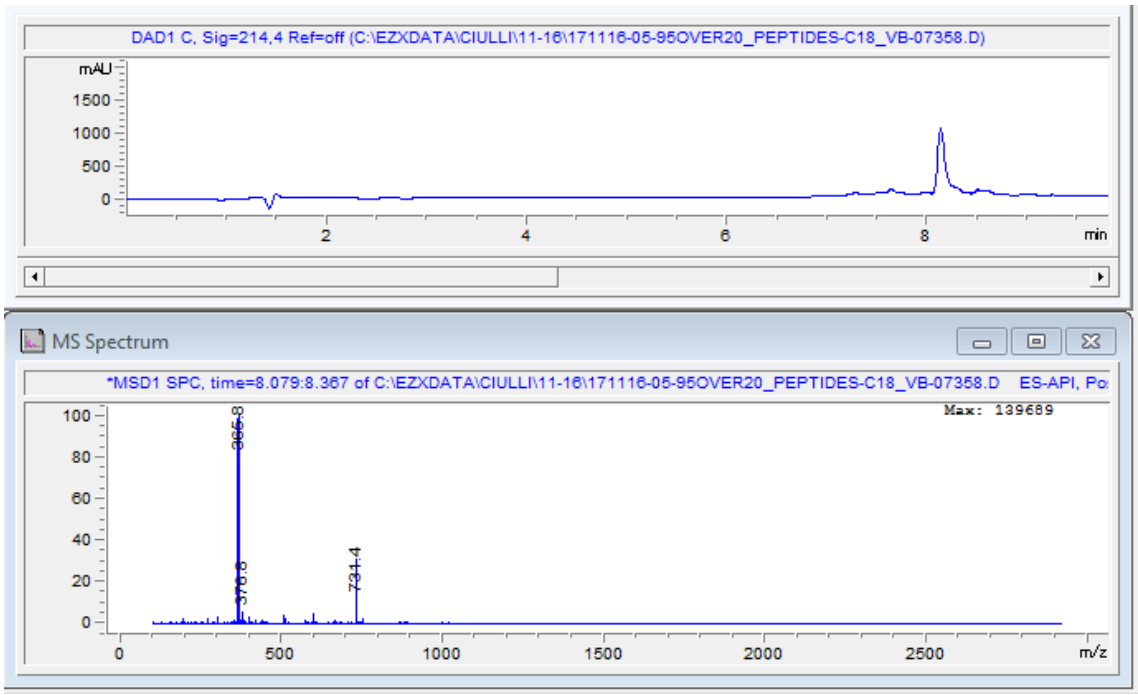


Figure 6.23 – Analytical LC-MS trace of peptide MSLKPR (UV trace recorded at 214 nm).

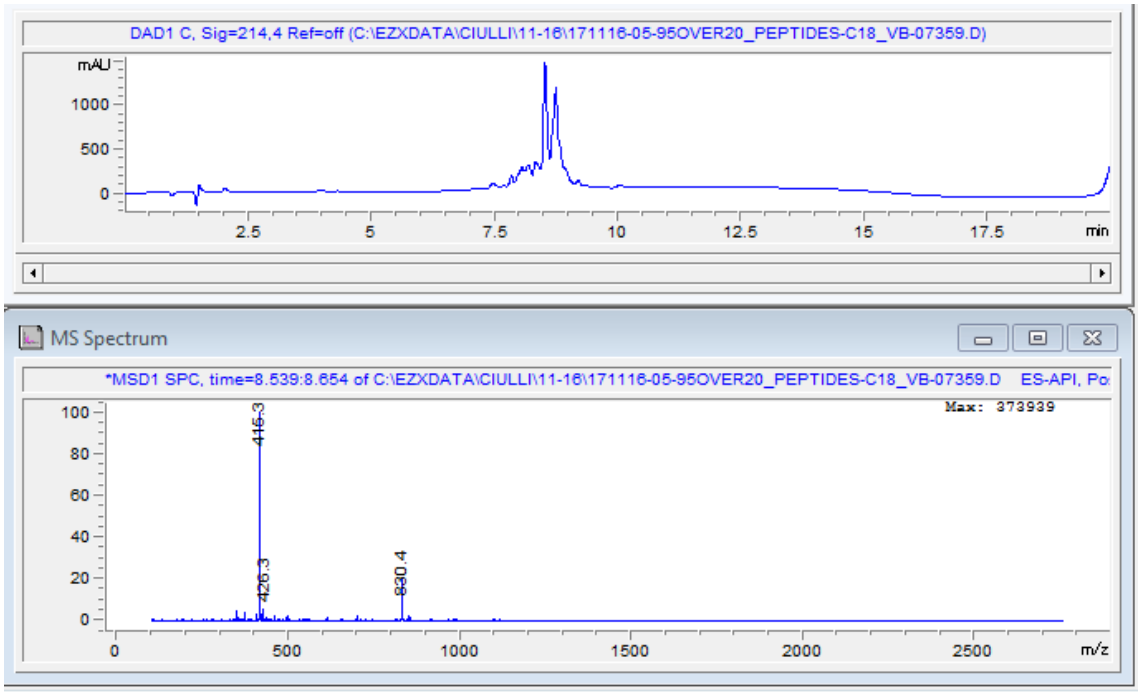


Figure 6.24 – Analytical LC-MS trace of peptide MSLKPRV (UV trace recorded at 214 nm).

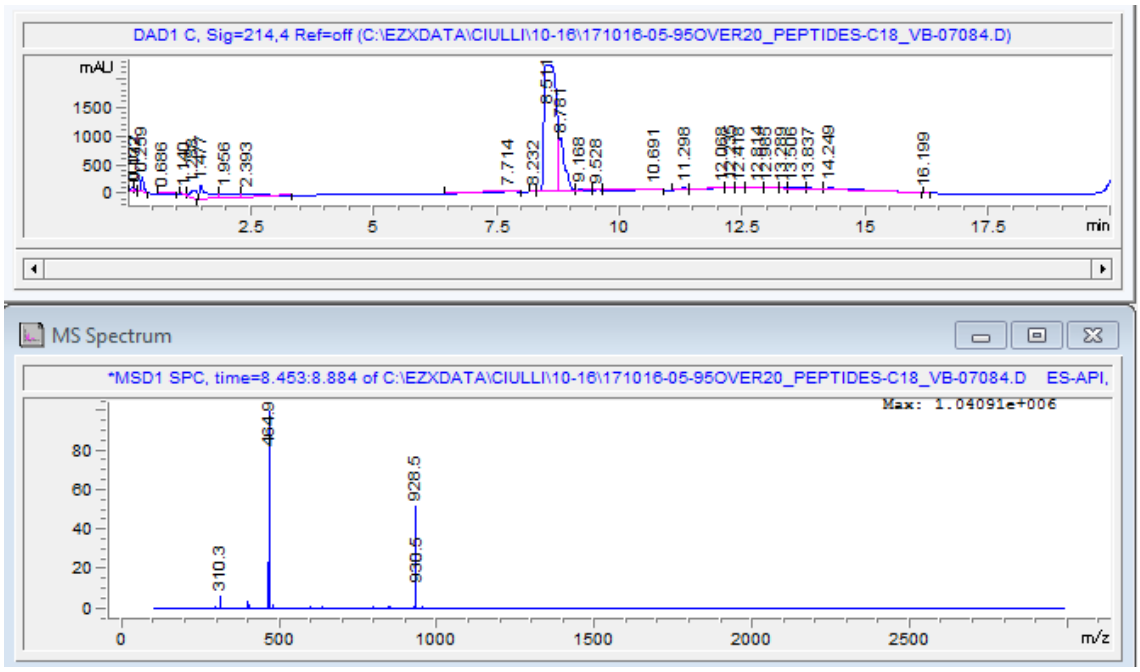


Figure 6.25 – Analytical LC-MS trace of peptide MSLKPRVV (peptide A) (UV trace recorded at 214 nm).

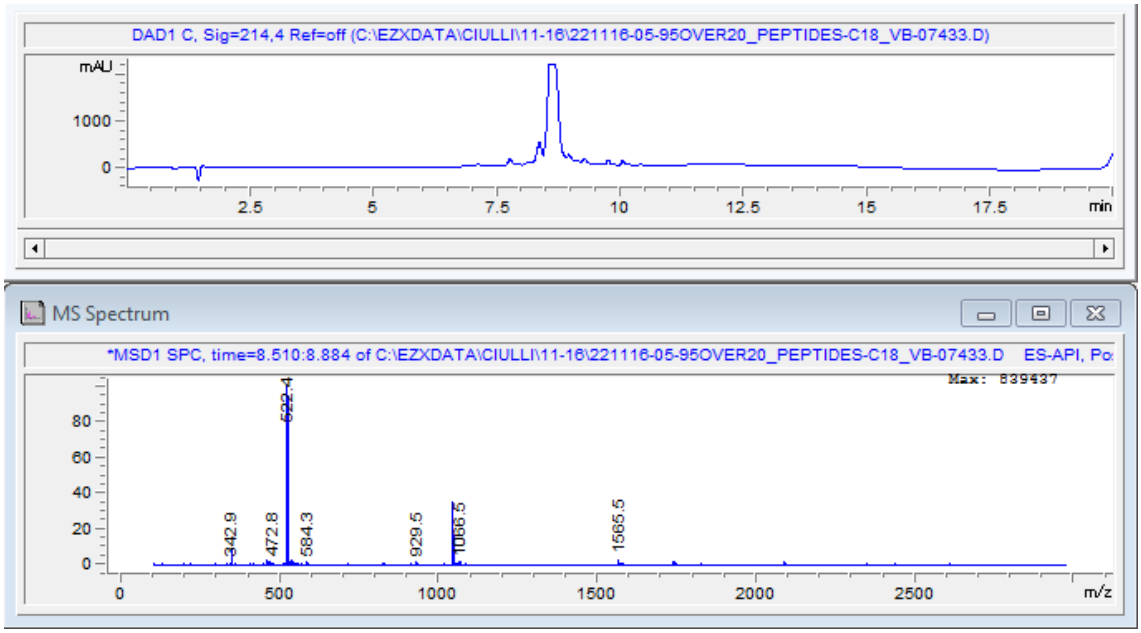


Figure 6.26 – Analytical LC-MS trace of peptide MSLKPRVVD (UV trace recorded at 214 nm).

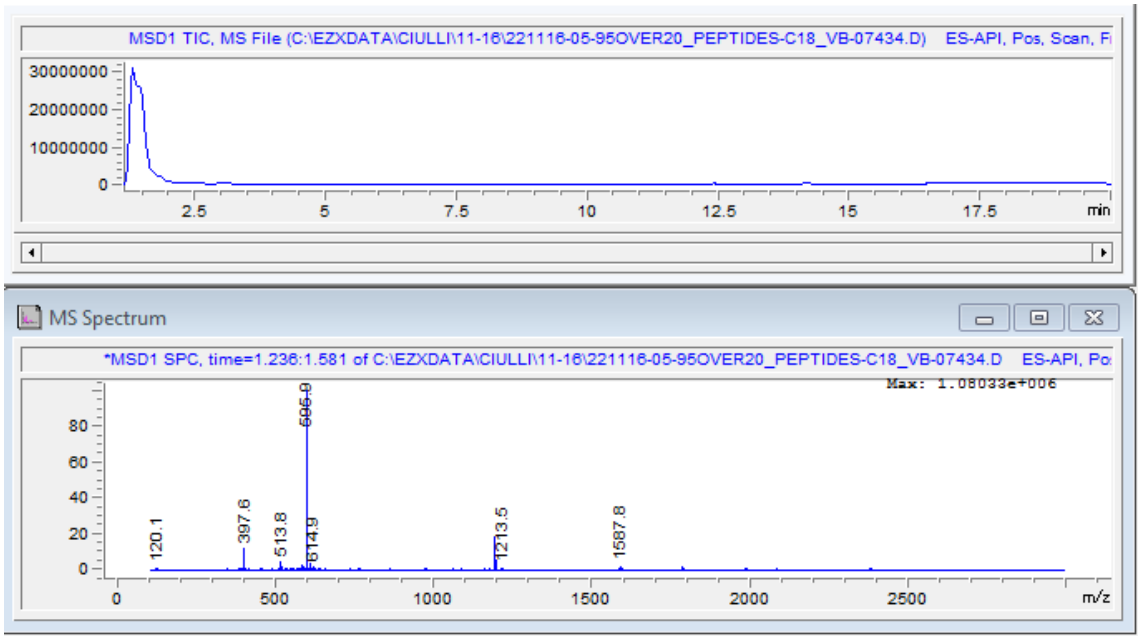


Figure 6.27 – Analytical LC-MS trace of peptide MSLKPRVVDF (UV trace recorded at 214 nm).

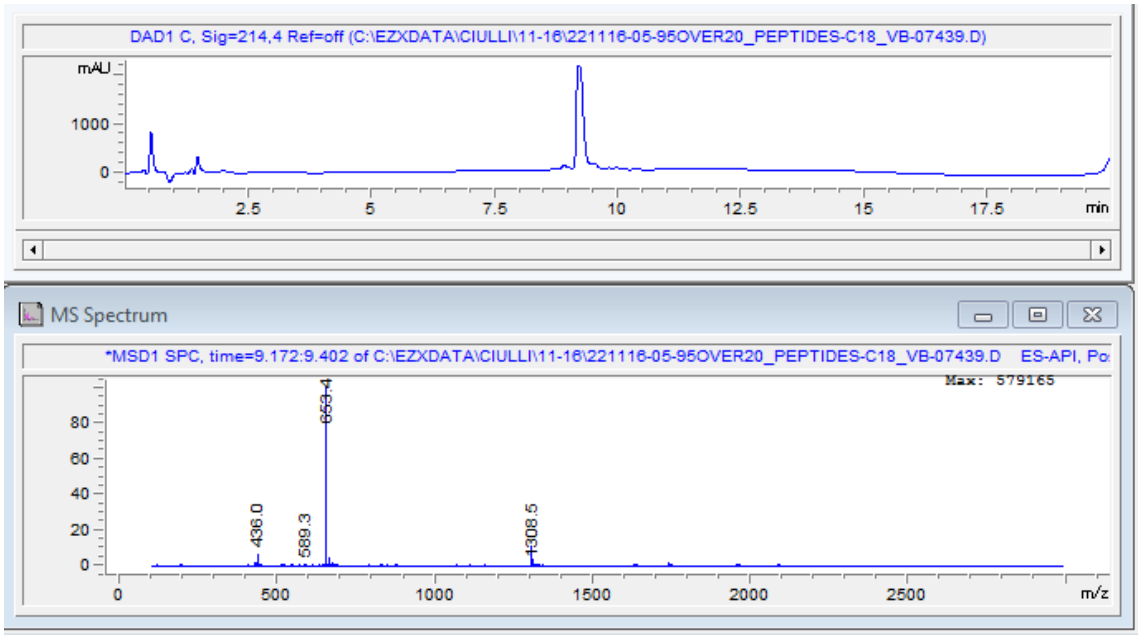


Figure 6.28 – Analytical LC-MS trace of peptide MSLKPRVVDFD (UV trace recorded at 214 nm).

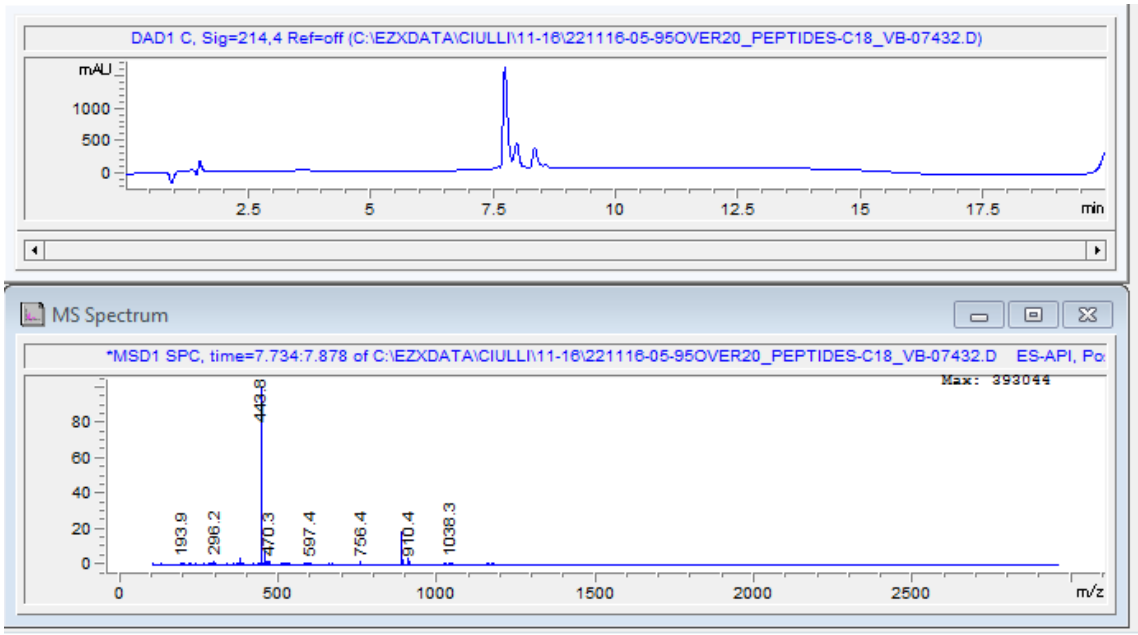


Figure 6.29 – Analytical LC-MS trace of peptide MSAKPRVV (UV trace recorded at 214 nm).

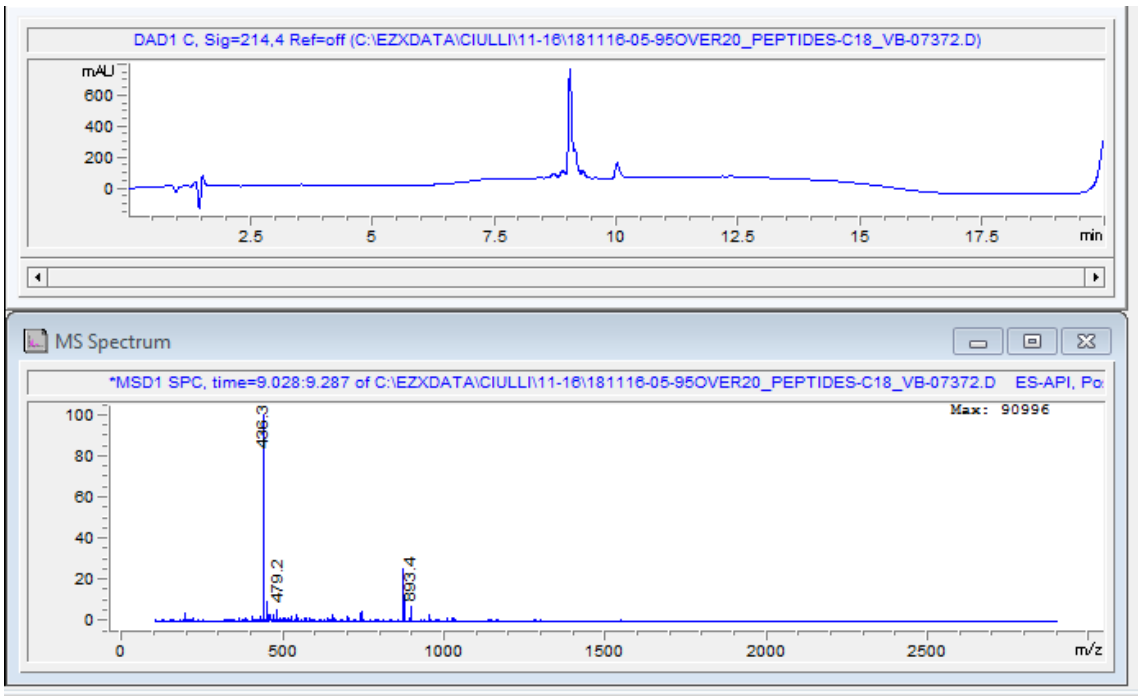


Figure 6.30 – Analytical LC-MS trace of peptide MSLAPRVV (UV trace recorded at 214 nm).

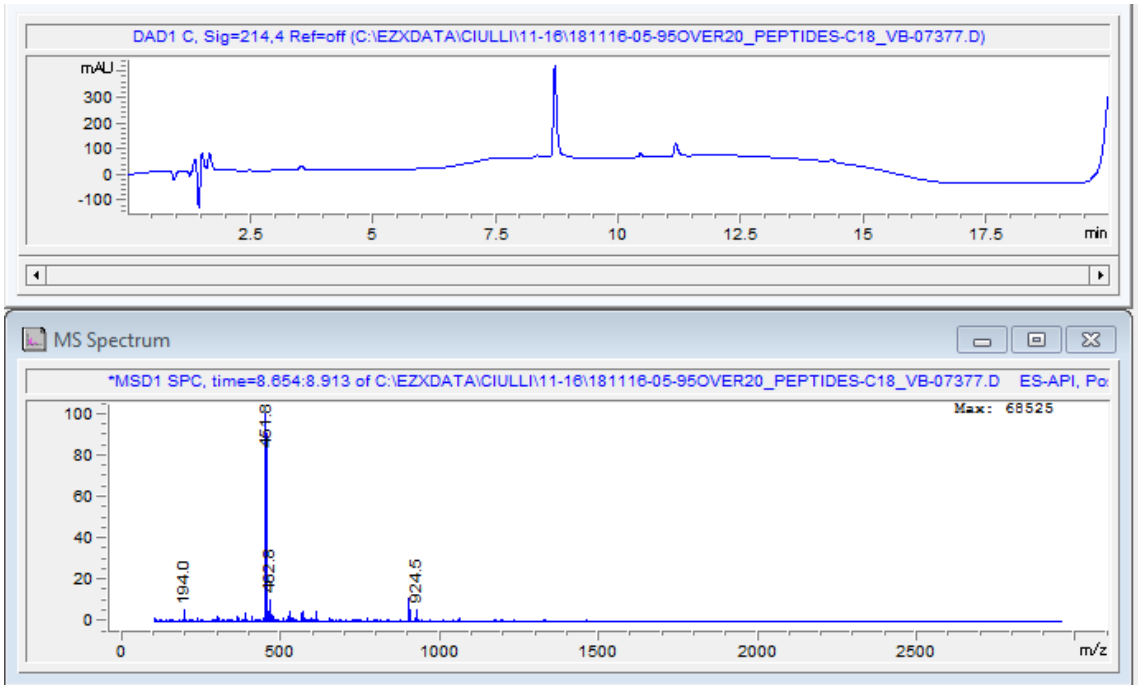


Figure 6.31 – Analytical LC-MS trace of peptide MSLKARVV (UV trace recorded at 214 nm).

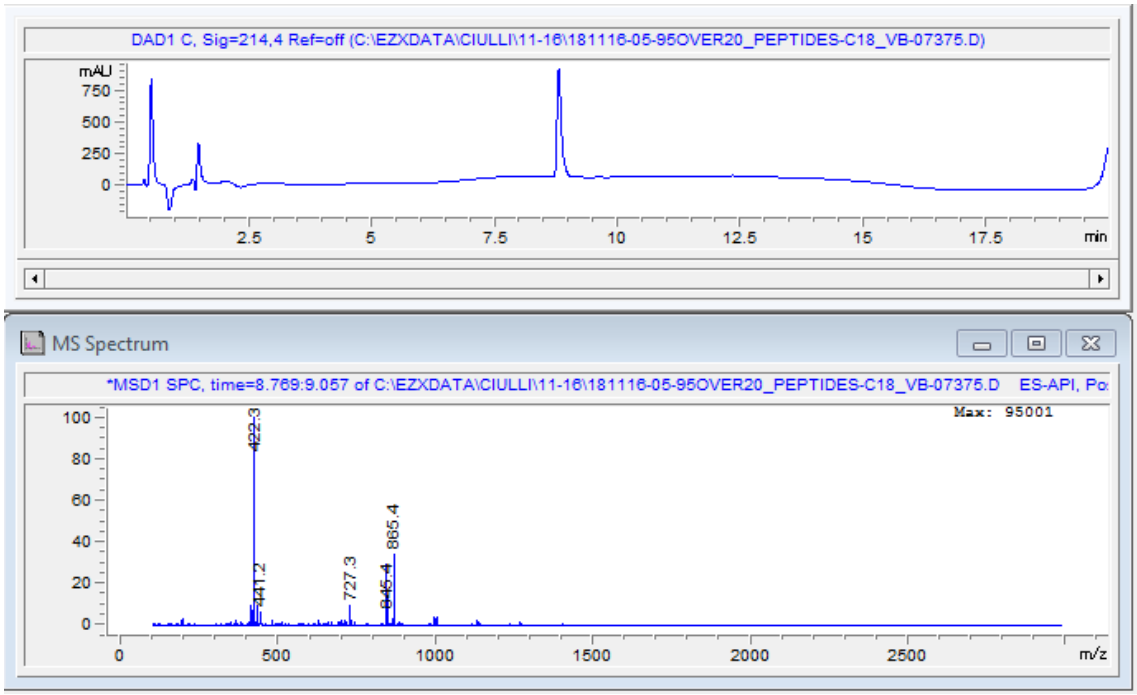


Figure 6.32 – Analytical LC-MS trace of peptide MSLKPAVV (UV trace recorded at 214 nm).

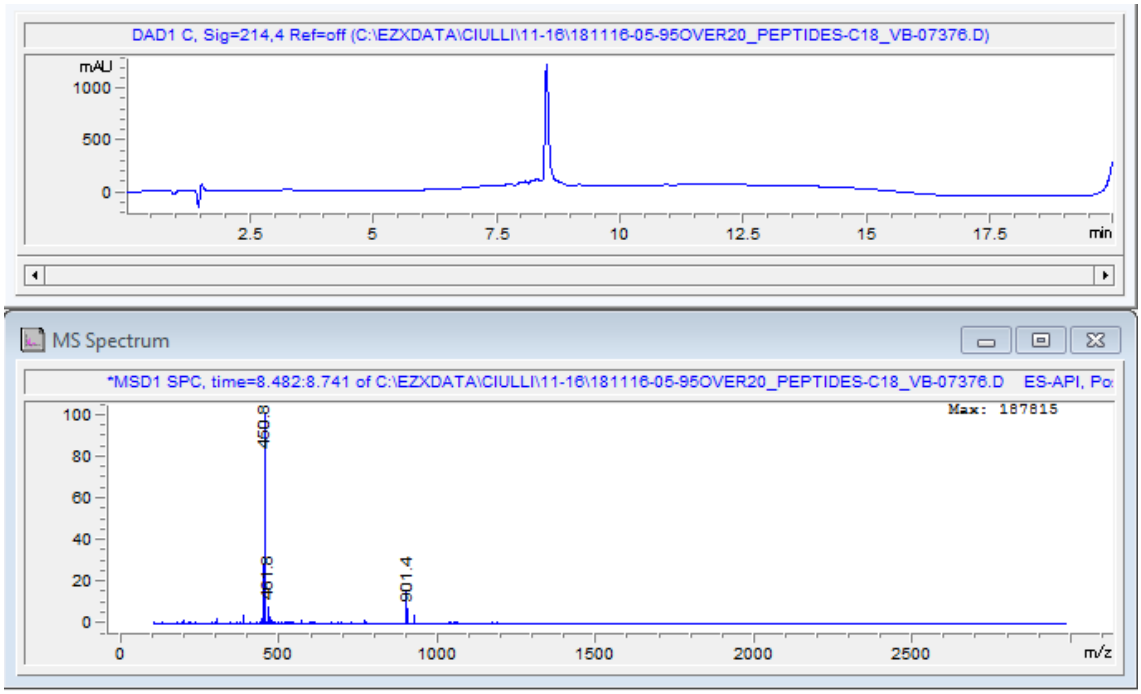


Figure 6.33 – Analytical LC-MS trace of peptide MSLKPRAV (UV trace recorded at 214 nm).

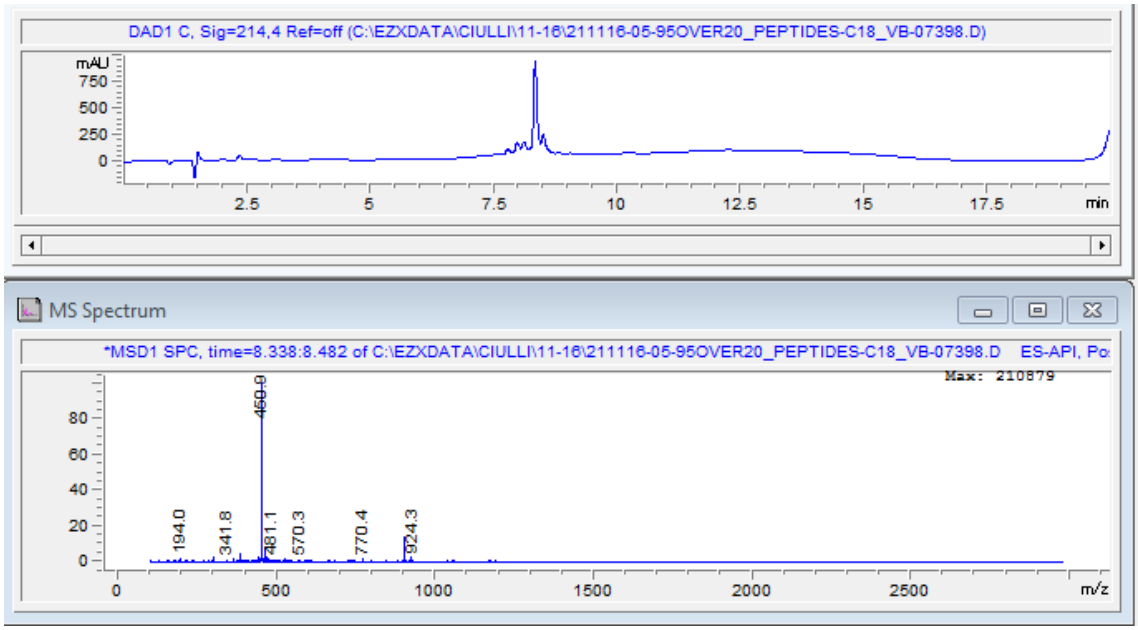


Figure 6.34 – Analytical LC-MS trace of peptide MSLKPRVA (UV trace recorded at 214 nm).

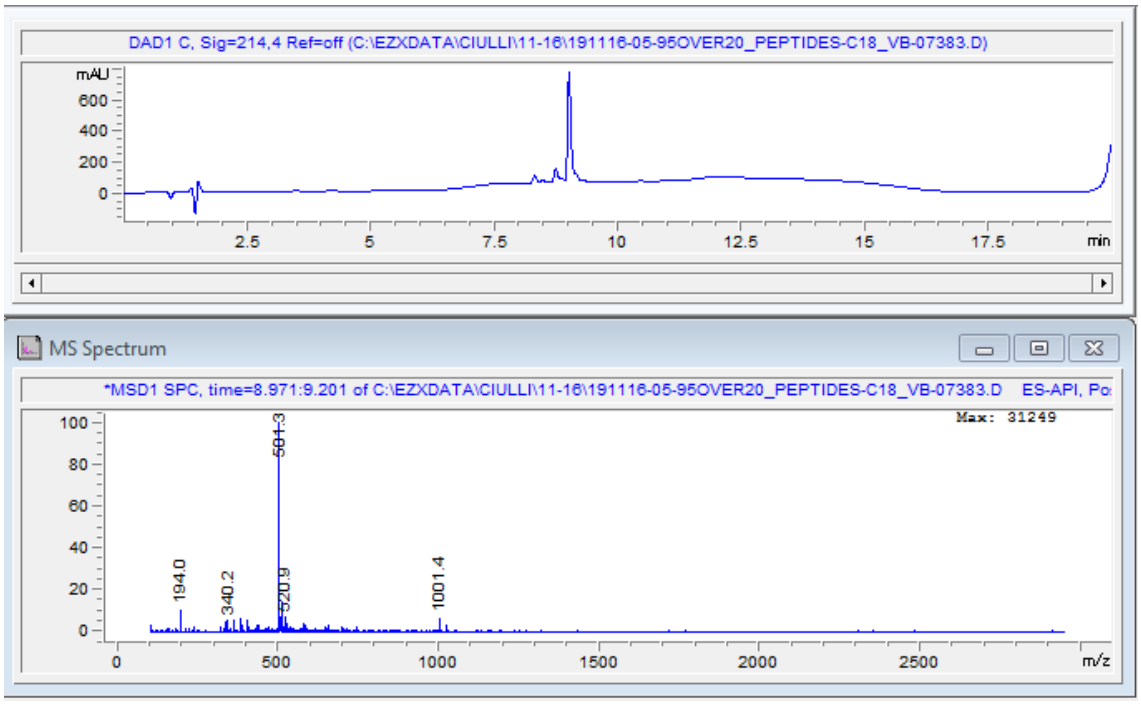


Figure 6.35 – Analytical LC-MS trace of peptide B (UV trace recorded at 214 nm).

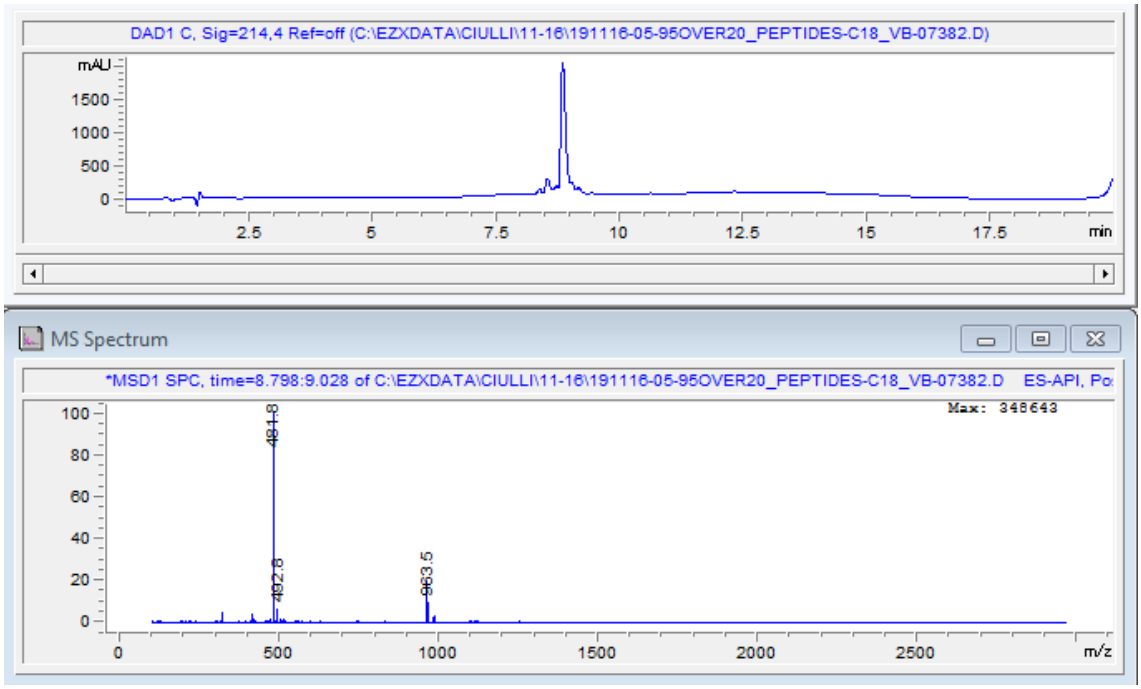


Figure 6.36 – Analytical LC-MS trace of peptide C (UV trace recorded at 214 nm).

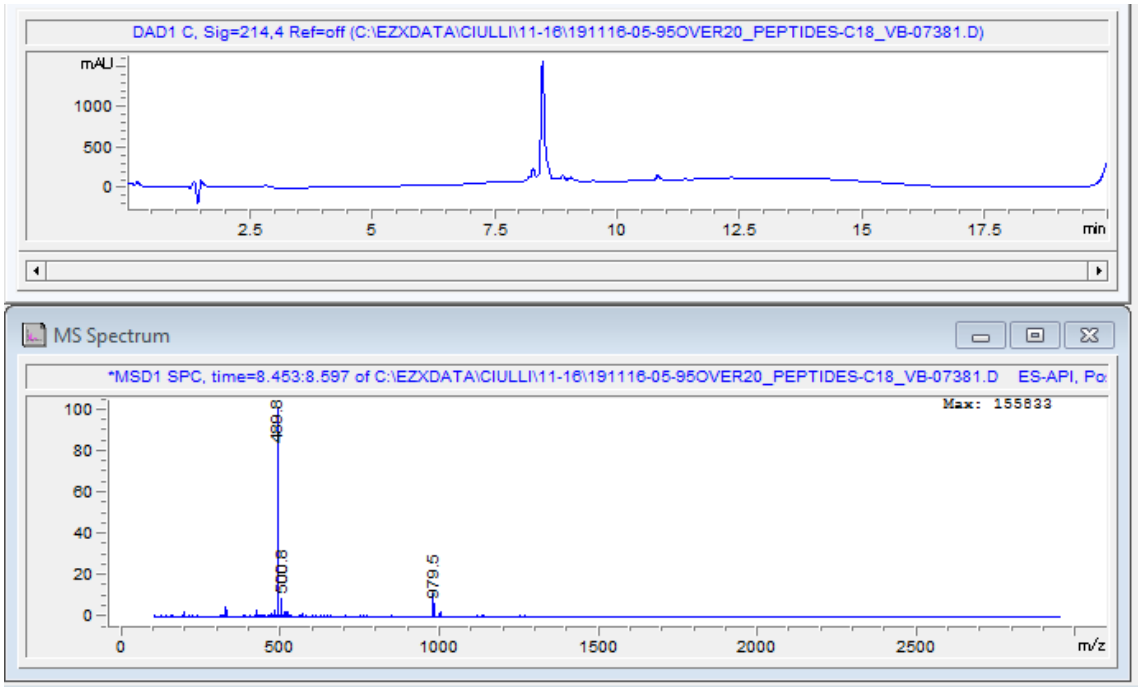


Figure 6.37 – Analytical LC-MS trace of peptide D (UV trace recorded at 214 nm).

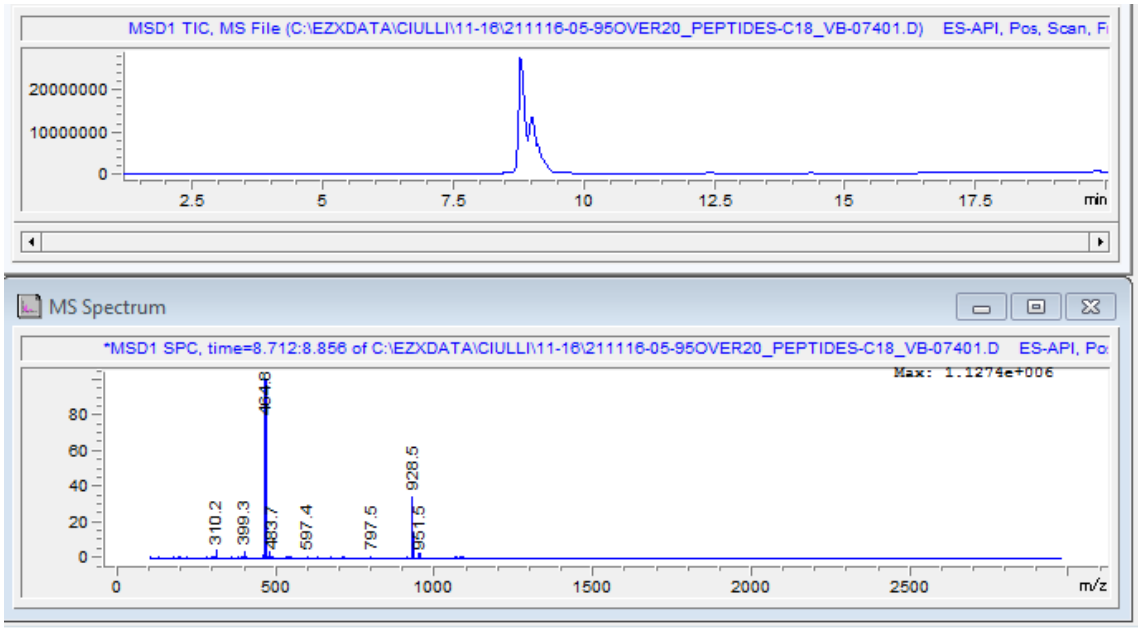


Figure 6.38 – Analytical LC-MS trace of peptide E (UV trace recorded at 214 nm).

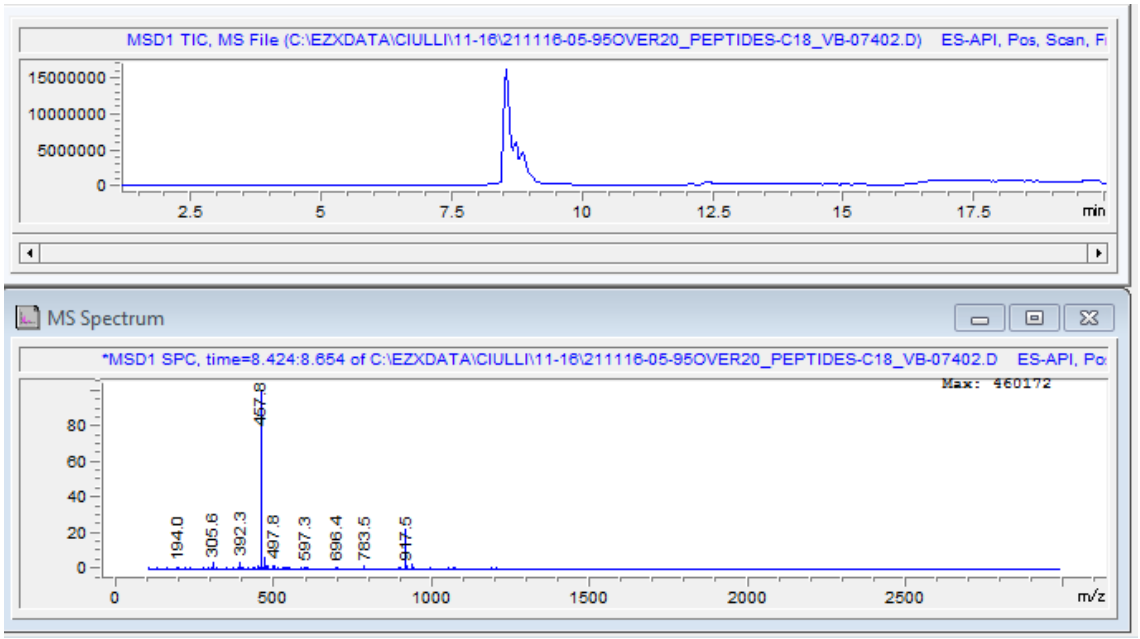


Figure 6.39 – Analytical LC-MS trace of peptide F (UV trace recorded at 214 nm).

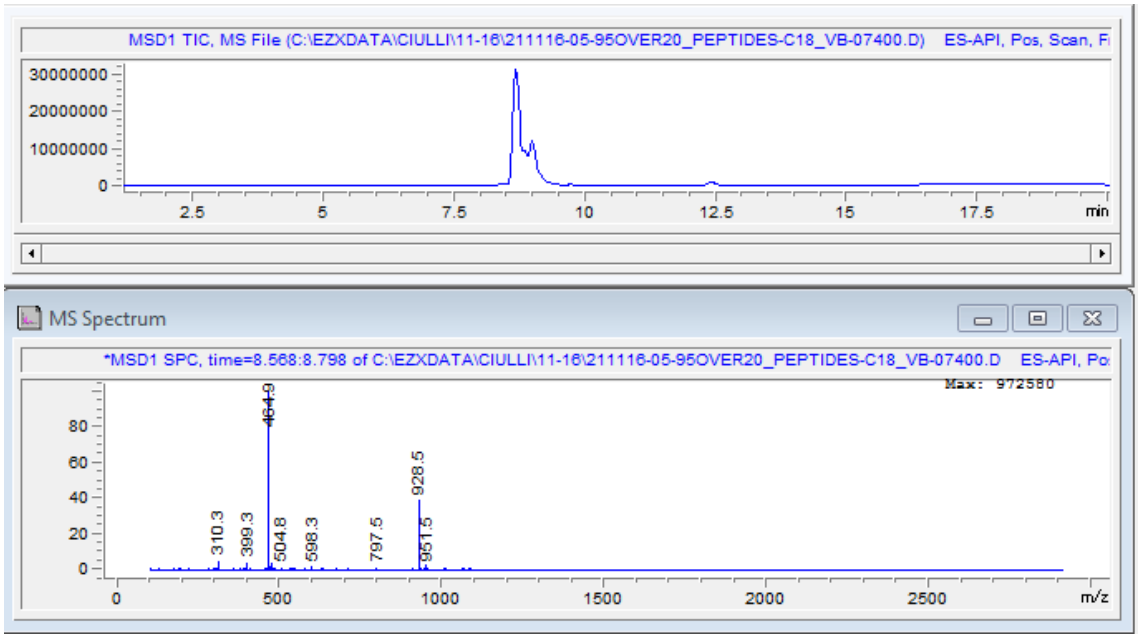


Figure 6.40 – Analytical LC-MS trace of peptide G (UV trace recorded at 214 nm).

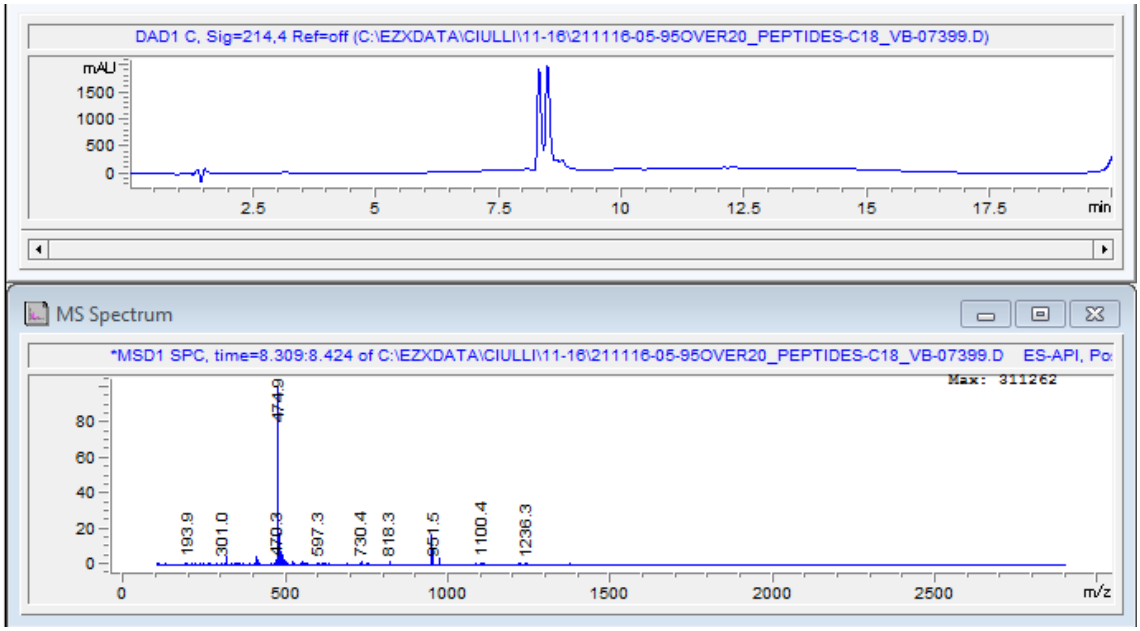


Figure 6.41 – Analytical LC-MS trace of peptide H (UV trace recorded at 214 nm).

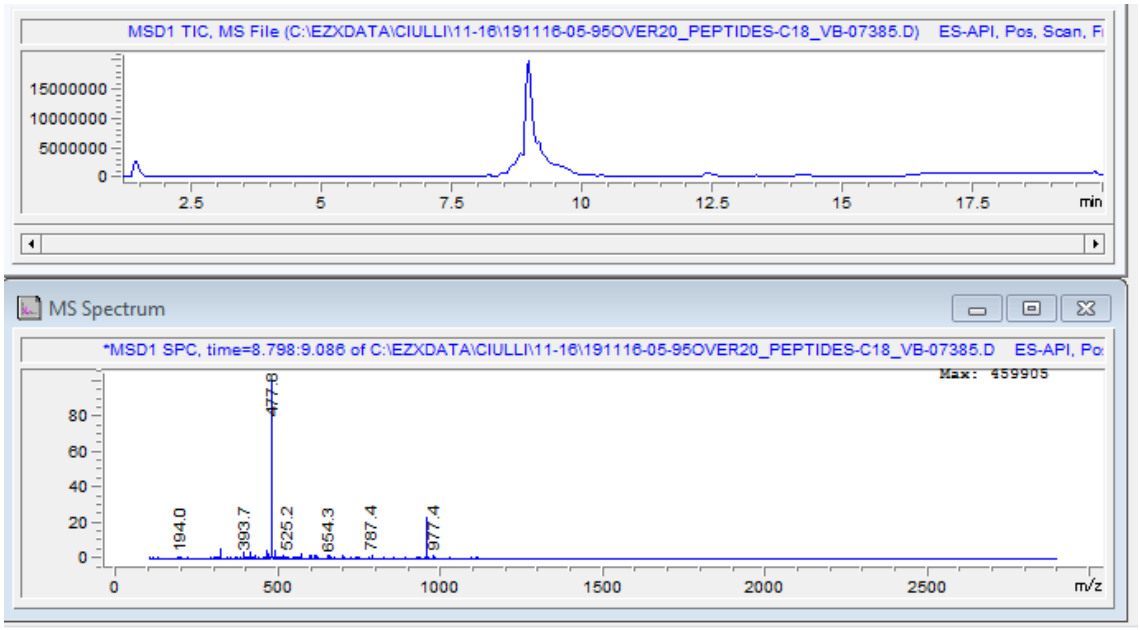


Figure 6.42 – Analytical LC-MS trace of peptide I (UV trace recorded at 214 nm).

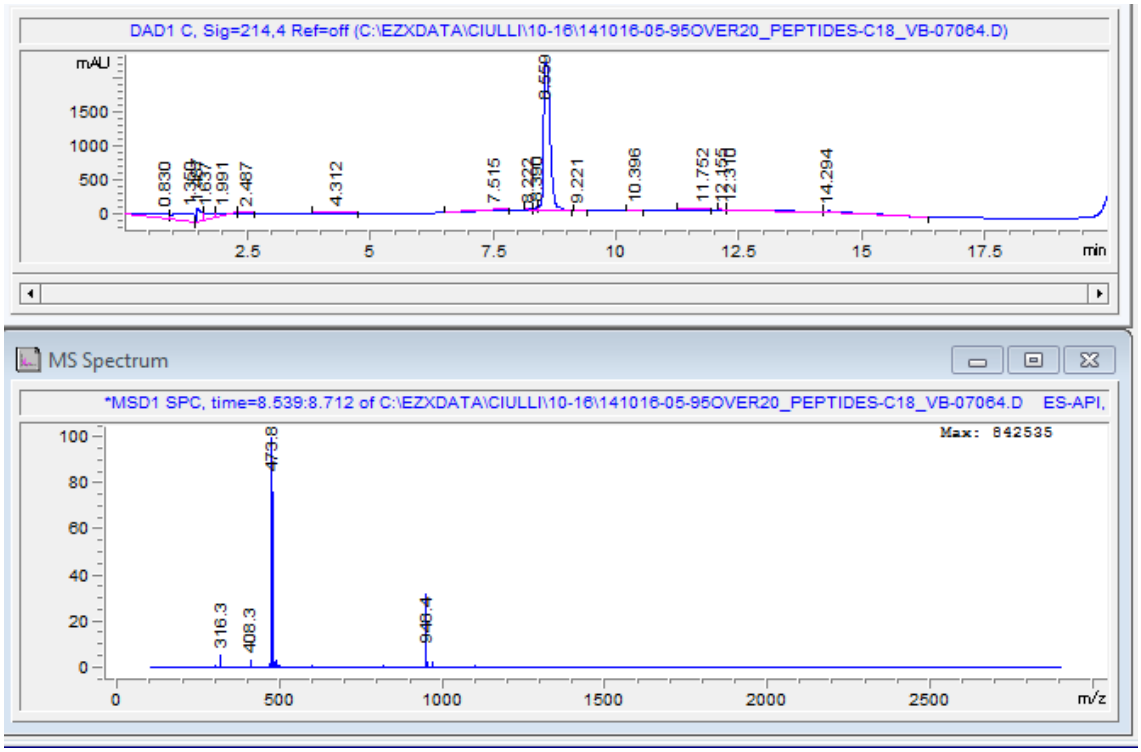


Figure 6.43 – Analytical LC-MS trace of peptide J (UV trace recorded at 214 nm).

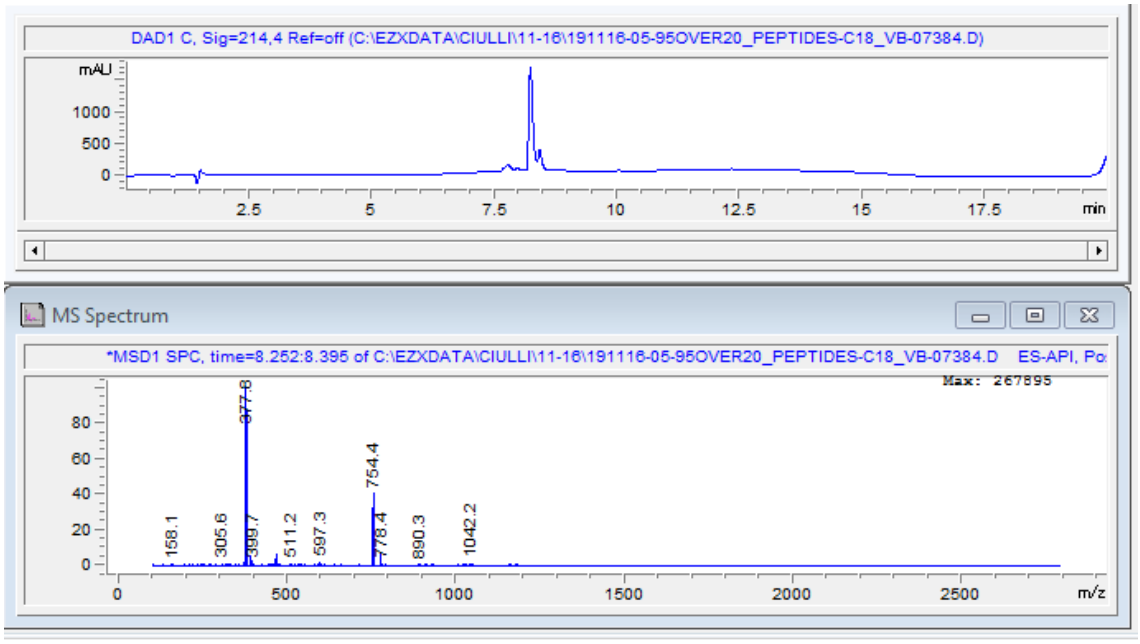


Figure 6.44 – Analytical LC-MS trace of peptide L (UV trace recorded at 214 nm).

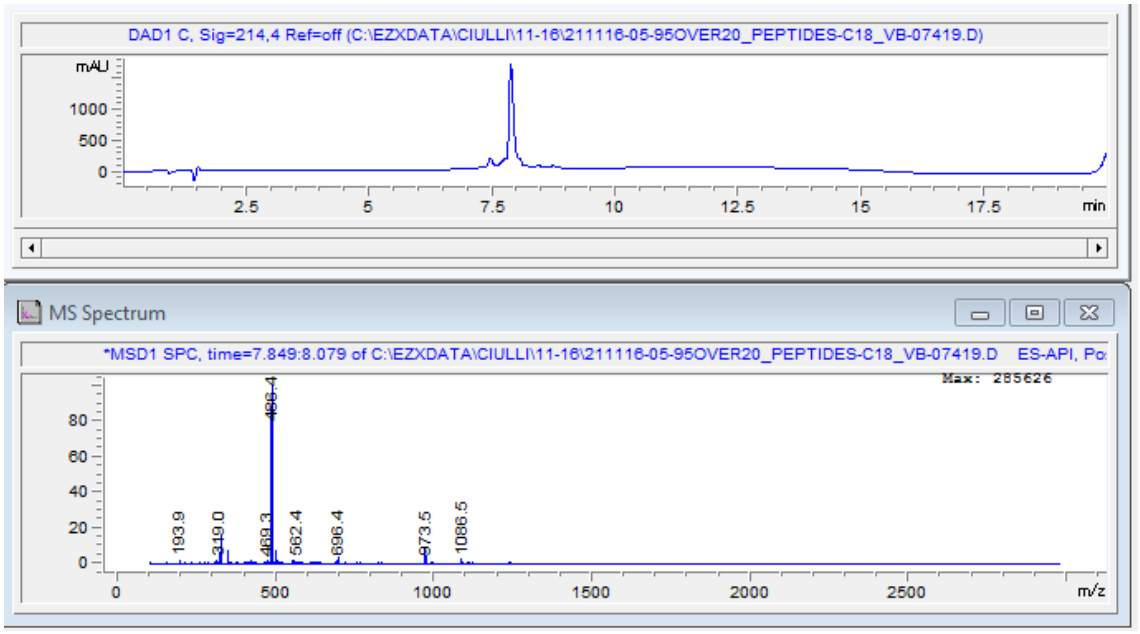


Figure 6.45 – Analytical LC-MS trace of peptide M (UV trace recorded at 214 nm).

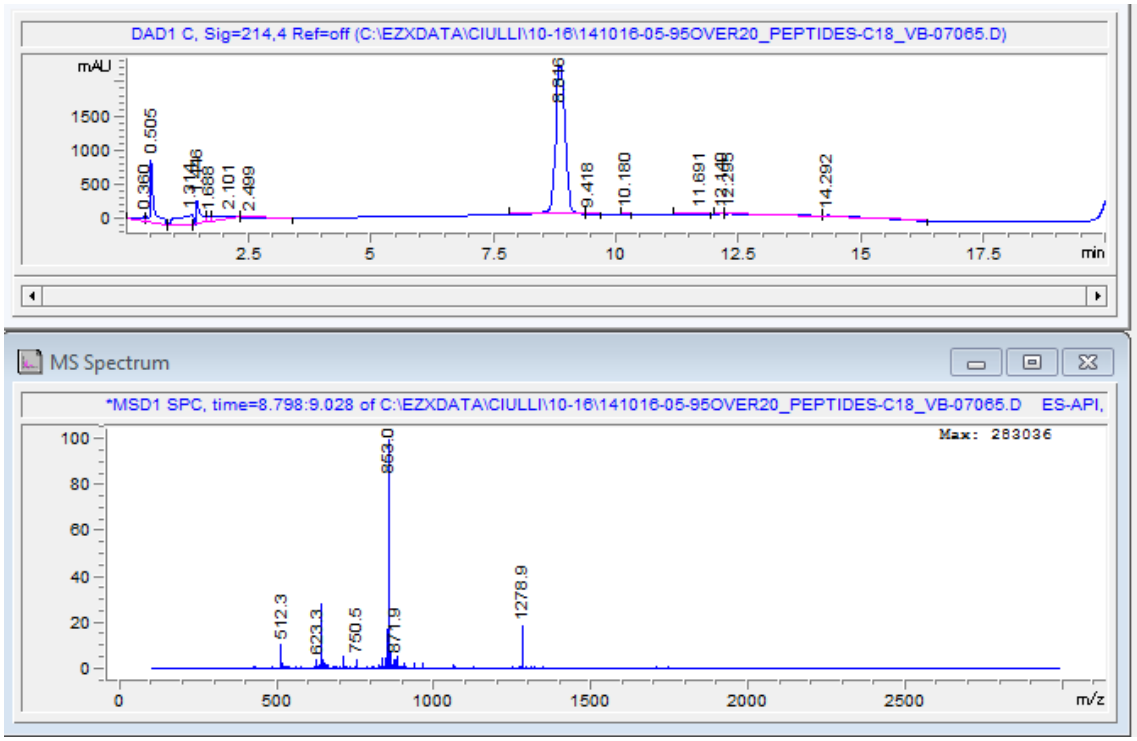


Figure 6.46 – Analytical LC-MS trace of Clamp peptide derived from peptide A (UV trace recorded at 214 nm).

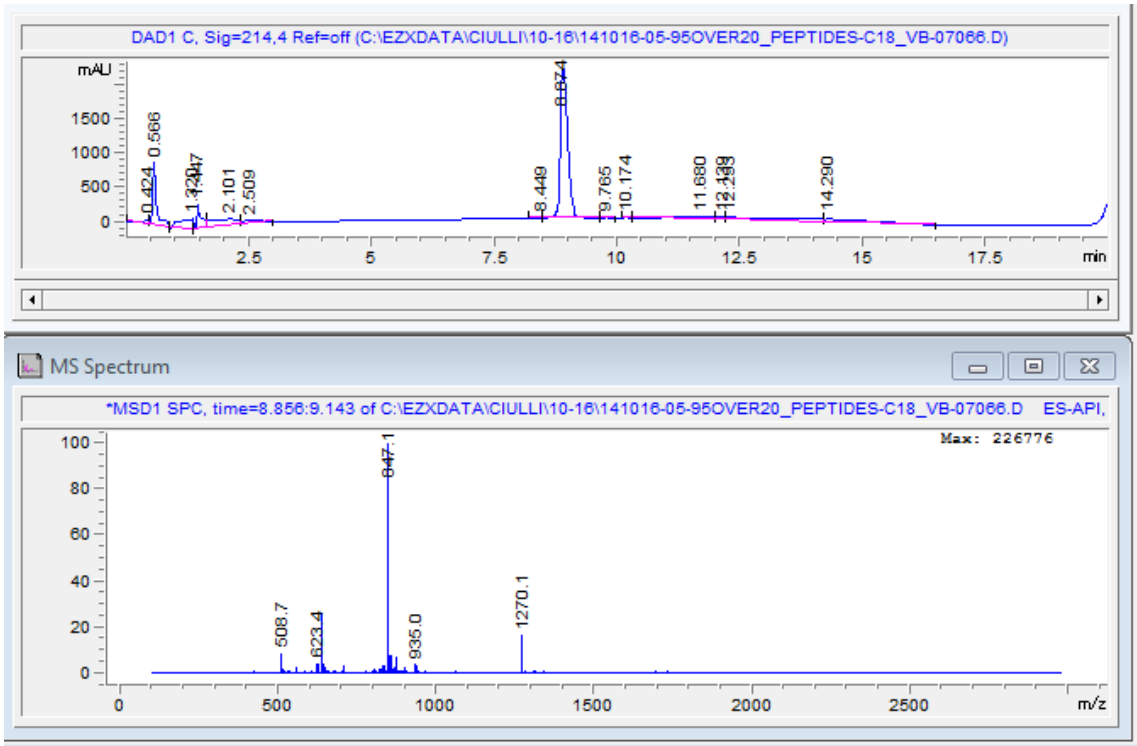


Figure 6.47 – Analytical LC-MS trace of Clamp peptide derived from peptide J (UV trace recorded at 214 nm).

6.4.2 Cyclic pentapeptides

The procedure for cyclisation was based on the protocol described by White *et al*.[139]

The linear peptides (0.03 mmole) were synthesised in the automated peptide synthesiser using a low loading capacity resin (Rink amide AM LL 100-200 mesh, Novabiochem, 0.2-0.4 mmole/g). For the side chain to backbone cyclisation, Fmoc-Asp(O-allyl) or Fmoc-Glu(O-allyl) were used as C-terminal residue. Next, the resin was transferred to a syringe filter and washed thoroughly with DMF, then DCM, and finally DMF again. The syringe was then capped and the resin was incubated with 1 eq. of $Pd(PPh_3)_4$ in THF and 10% piperidine for 3 h under constant shaking. After this step, the resin was washed again with DMF and then with a mixture of 5% (w/v) NaDEDTC and 5% (v/v) DIPEA in DMF and finally with DMF again. The cyclisation was performed with 3 eq. of HATU, 3.2 eq. of HOAT and 5 eq. of DIPEA in DMF for 3 h, shaking. Then the resin was washed with DCM and the peptides were cleaved from the resin with 90% TFA in water for another 3 h with constant shaking. Finally the resin was filtered and the cyclic peptides were precipitated in cold diethyl ether. The cyclic pentapeptides were dissolved in water and purified on a Gilson Preparative HPLC System with a Waters X-Bridge C18 column (100 mm x 19 mm; 5 µm particle size) and a gradient from 5% to 95% of acetonitrile in water over 20 min, flow 25 mL/min, with 0.1% ammonia in the aqueous phase. The identity and purity of the peptides was confirmed by analytical HPLC-MS on an Agilent Technologies 1200 series HPLC connected to an Agilent Technologies 6130 quadrupole LC/MS linked to an Agilent diode array detector (Figure 6.48 – 6.53). The column used was an Agilent ZORBAX StableBond 80Å C18, 4.6 x

250 mm, 5 µm column and the compounds were eluted with a gradient of 5−95% acetonitrile/water + 0.1% ammonia.

6.4.3 Bicyclic peptides

The linear peptides were synthesised according to the method described in section 6.4. For the cyclisation reaction, the protocol consisted of an alkylation reaction achieved by the dilution of the peptide to 0.5 mM in a solution of 20% ACN and 80% 40 mM NH_4HCO_3 pH 8.0 in the presence of 1 mM of TBMB for 1 h at 30 °C.[153] At the end of the reaction, the peptides were lyophilised in the centrifugal evaporator (Genevac EZ-2 series, SP Scientific). The crude peptides were dissolved in solution of H_2O:ACN at different ratios according to the optimal condition for each peptide and purified on a Gilson Preparative HPLC System with a ZORBAX Stable bond 80Å C18 column (9.4 mm x 250 mm; 5 µm particle size) and a gradient from 5% to 95% of acetonitrile in water over 20 min, flow 25 ml/min, with 0.1% TFA in the aqueous phase. The fractions containing the bicyclic peptides were lyophilised and the identity and purity of the peptides was confirmed by analytical HPLC-MS on an Agilent Technologies 1200 series HPLC connected to an Agilent Technologies 6130 quadrupole LC/MS, connected to an Agilent diode array detector. The column used was an Agilent ZORBAX StableBond 80Å C18, 4.6 mm x 250 mm, 5 µm column and the compounds were eluted with a gradient of 5−95% acetonitrile/water + 0.1% TFA.

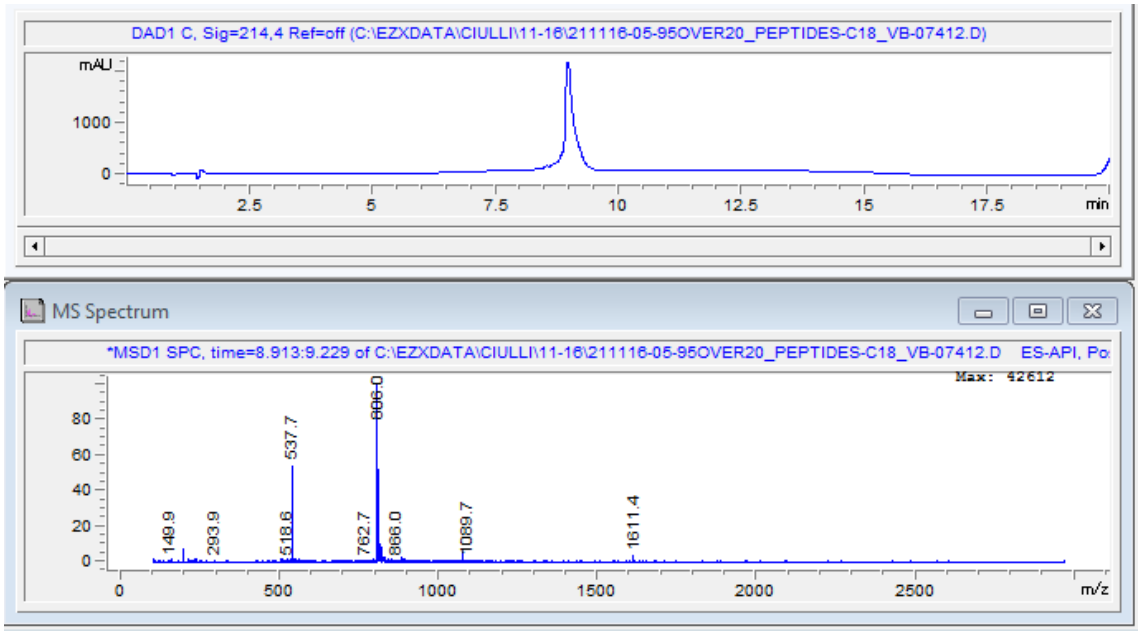


Figure 6.48 – Analytical LC-MS trace of Peptide 34 (UV trace recorded at 214 nm).

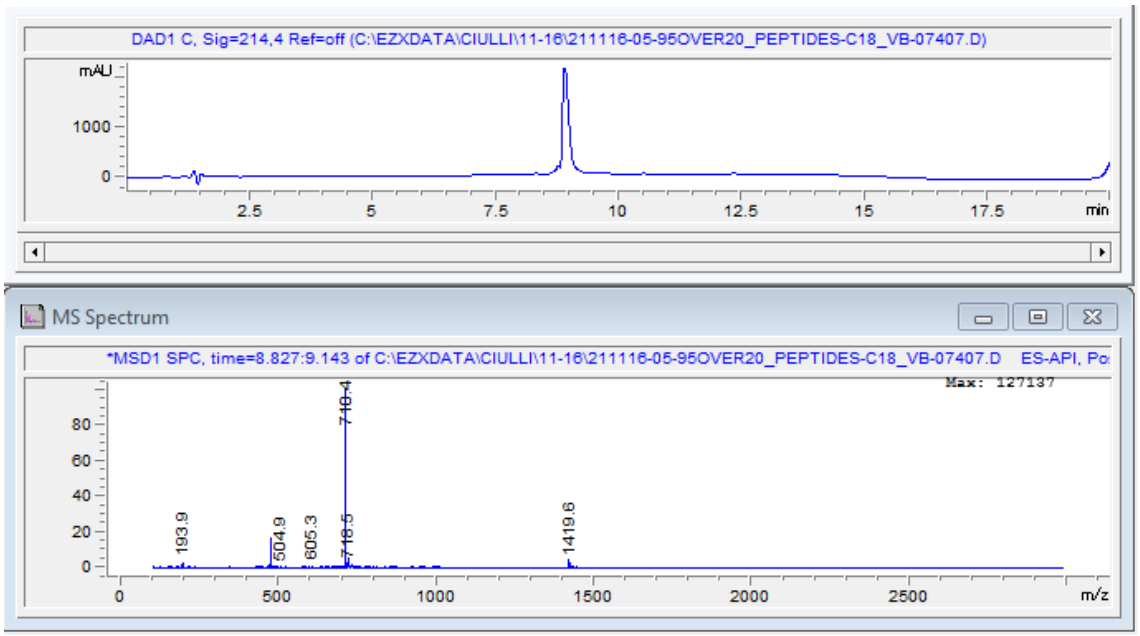


Figure 6.49 – Analytical LC-MS trace of Peptide 66 (UV trace recorded at 214 nm).

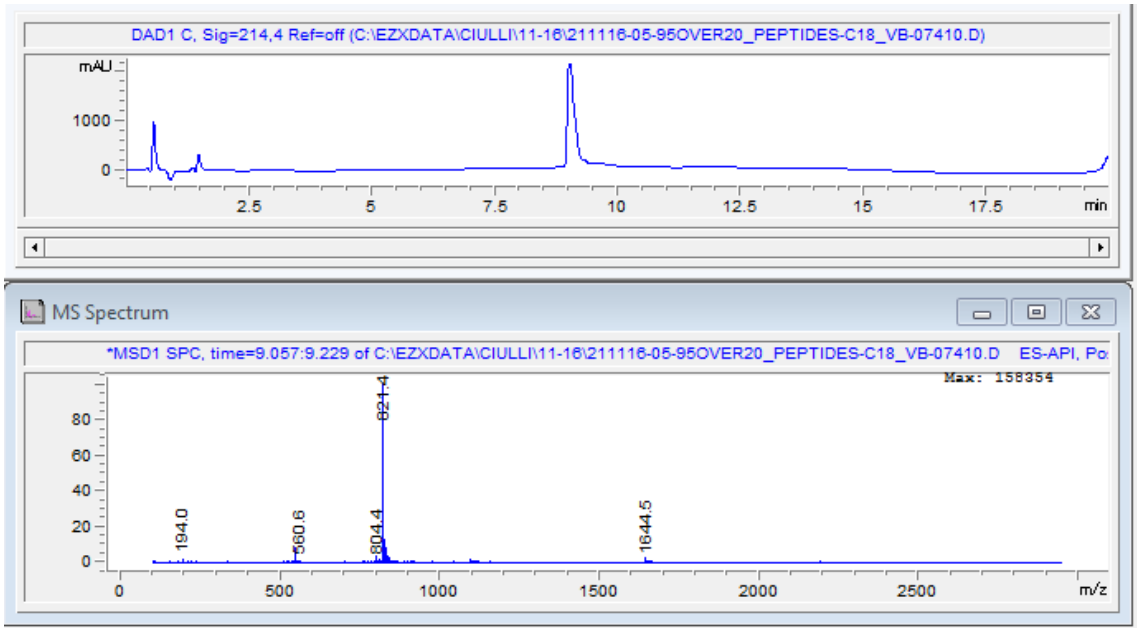


Figure 6.50 – Analytical LC-MS trace of Peptide 86 (UV trace recorded at 214 nm).

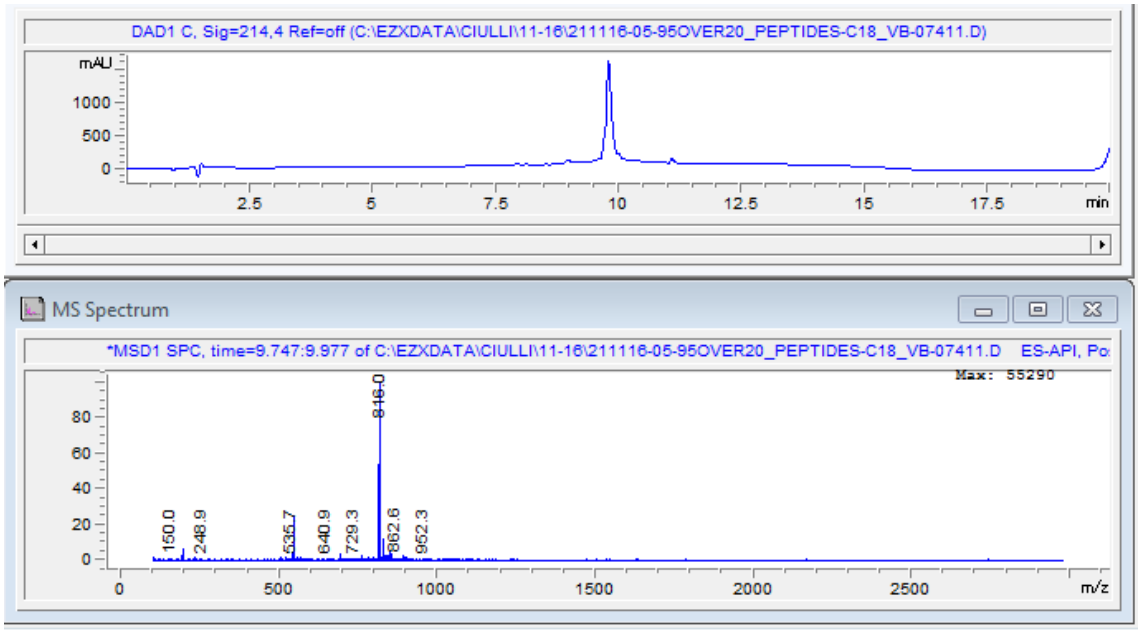


Figure 6.51 – Analytical LC-MS trace of Peptide 109 (UV trace recorded at 214 nm).

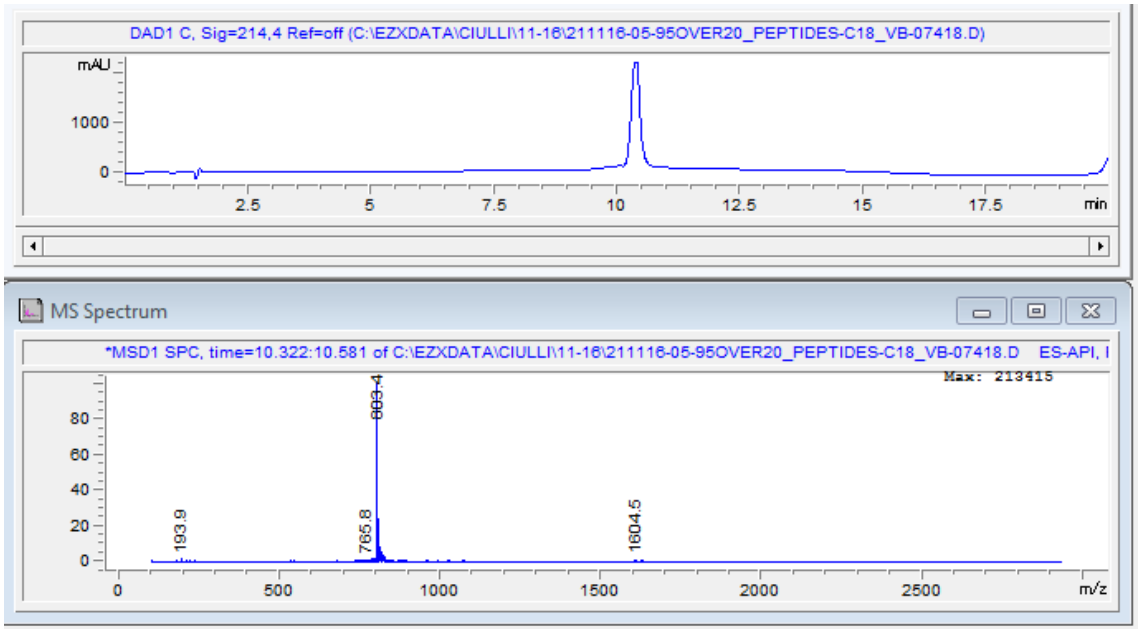


Figure 6.52 – Analytical LC-MS trace of Peptide 113 (UV trace recorded at 214 nm).

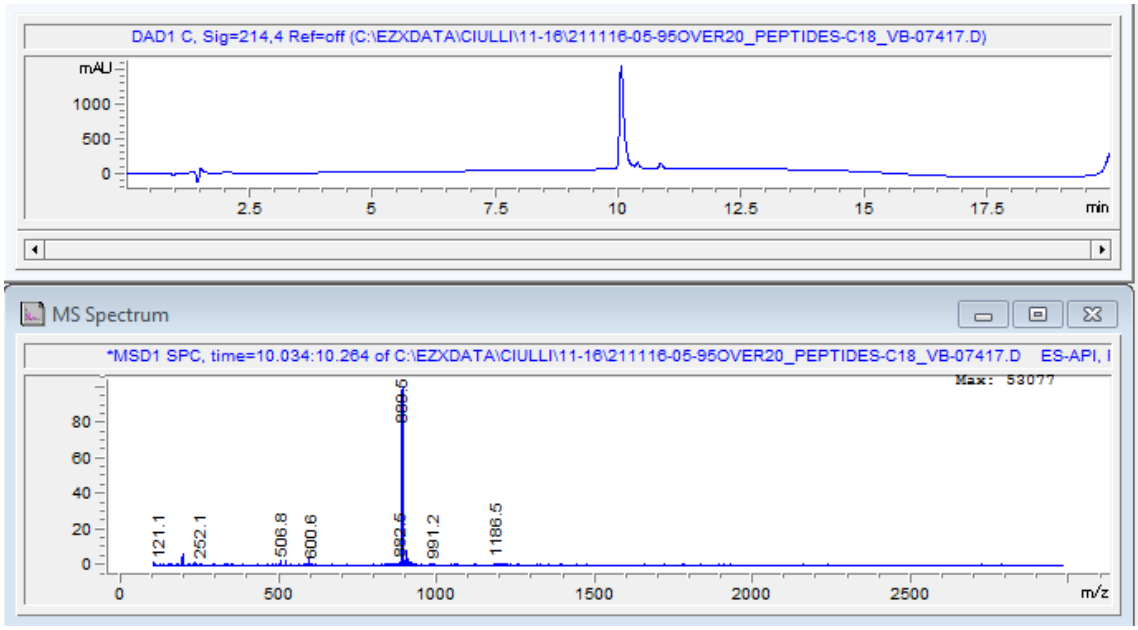


Figure 6.53 – Analytical LC-MS trace of Peptide 115 (UV trace recorded at 214 nm).

6.5 Biophysical techniques

6.5.1 Isothermal titration calorimetry

ITC experiments were carried out in an ITC200 microcalorimeter (Malvern). All protein solutions were dialysed into 100 mM Bis-tris propane pH 8.0, 50 mM NaCl, 2 mM TCEP prior to the titrations.

For the titrations of protein into protein, VBC (wild type and variant proteins) or SBC (wild type and variant proteins) were titrated into Rbx1–Cul2. For the experiments with $Cul5_{NTD}$ the latter was loaded in the syringe and titrated against VBC (wild type and variant proteins) or SBC (wild type and variant proteins). The titrations consisted of 19 injections of 2 µl each (120 sec spacing and 600 rpm stirring speed) at 25 °C or 30 °C.

For the titrations of peptide into protein, the peptide (< 1 mM) was titrated into VBC (100 µM unless noted). The titrations consisted of 19 injections of 2 µl each (120 sec spacing and 600 rpm stirring speed) at 30 °C.

Control experiments were performed injecting the titrant into buffer and were subtracted to the relevant experiment to account for heat of dilution. Data analyses for the ITC experiments were performed using the MicroCal Origin 7.0 software package. Binding enthalpy, dissociation constants, and stoichiometry were determined by fitting the data using a one-set-of-sites binding model. The experiments were performed twice for to assess data reproducibility.

6.5.2 Differential scanning fluorimetry

Differential scanning fluorimetry (DSF) assays were performed in a Mx3005P Real Time PCR machine (Stratagene) or in a CFX96 Touch Real-time PCR detection system (Biorad). SYPRO Orange (Invitrogen Molecular Probes) at a dilution of 1:1000 was used as a reporter dye to monitor the thermal

denaturation process of the proteins. Samples were assayed on a 96-well plate with final protein concentration of 7.5 µM. The temperature was raised from 25 °C to 95 °C in increments of 1 °C per minute, collecting fluorescence readings at the end of each interval. Each sample was run in triplicate and data were reported as an averaged value.

6.5.3 Biolayer interferometry

BioLayer Interferometry experiments were performed in an Octet RED384 (FòrteBio). The biotinylated proteins (25 µg/ml) were immobilised on Super Streptavidin biosensor (FòrteBio) and the reference sensors were quenched with biocytin (1 mg/ml). The assays were conducted at 25 °C, in 20 mM HEPES pH 7.6, 100 mM NaCl, 1 mM DTT and 0.02% (v/v) Tween-20 buffer. The experiments were performed in triplicate in a black solid 384-well plate under agitation at 1000 rpm. Common cycle steps for the analyses included 60 sec of baseline equilibration in buffer, 60 sec for association in wells containing the ligand and 180 sec for the dissociation in wells containing only buffer. The response of the reference tips was subtracted from the signal to account for unspecific binding. The data points were fitted using a 1:1 model and the K_d values were calculated according to the following equations[80]:

Association phase:

$$y = R_{max}\frac{1}{1+\frac{k_d}{k_a * [Analyte]}}(1-e^{-(k_a*[Analyte]+k_d)x})$$

Dissociation phase:

$$y = y_0 e^{-k_d (x-x_0)}$$

$$y_0 = R_{max}\frac{1}{1+\frac{k_d}{k_a * [Analyte]}}(1-e^{-(k_a*[Analyte]+k_d)x_0})$$

.

6.5.4 NMR spectroscopy

NMR experiments were carried out in an AV-500 MHz Bruker spectrometer equipped with a 5 mm CTPXI ^{1}H-$^{13}C/^{15}N$/D Z-GRD cryoprobe.

For ligand-observed experiments three samples were prepared:

A) 30 μM VBC (protein control)

B) 500 μM-1.5 mM peptide (ligand control)

C) 30 μM VBC + 500 μM-1.5 mM peptide

The samples were prepared in 3-mm capillary tubes (Bruker) using a buffer containing 20 mM KH_2PO_4 pH 7.0, 50 mM KCl, 1 mM DTT, 0.02% NaN_3, 20 μM of TSP (for calibration purposes) and 15 % D_2O in a total volume of 200 μl. Samples B and C were prepared from the same ligand solution to avoid differences in concentration and while protein was added to C, the same amount of buffer was added to B. The experiments were carried out at 20 °C using the CPMG[78] sequence previously developed and optimised in the Ciulli laboratory. The resulting spectra were analysed in TopSpin.

For protein-observed experiments were used 100 μM $^{2}H,^{15}N$-VBC samples in a buffer containing 20 mM KH_2PO_4 pH 7.0, 50 mM KCl, 1 mM DTT, 0.02% NaN_3 and 15% of D_2O. The total volume of the sample was 200 μl and the experiments were performed in a capillary tube. The 2D $^{1}H,^{15}N$-HSQC-TROSY spectra (in the presence or absence of peptides) were recorded with 32 scans and acquisition times of 200 ms for ^{1}H and 100 ms for ^{15}N, at 30 °C. The spectra were analysed in CCPNMR[174] and the chemical shift perturbation (CSP) were calculated according to the following equation: $CSP = \sqrt{\Delta H^2 + (\Delta N \times 0.14)^2}$, where ΔH is the change in proton chemical shift, ΔN is the change in nitrogen chemical shift and 0.14 is a scaling factor required to account for the difference in the range of amide proton and amide nitrogen chemical shifts.[79] The backbone assignment of VBC has been made available as by Dr. Mark Bycroft (Laboratory of Molecular Biology, MRC, Cambridge) and shared as a gift.

6.5.5 AlphaLISA

For the AlphaLISA experiments Anti-6xHis acceptor beads and Streptavidin donor beads (PerkinElmer) were used. The competition assay was performed in a 384-well plate by mixing $V_{6xHis}BC$ (500 nM) and biotinylated Rbx1–Cul2 (150 nM) and the competitor in a concentration range from 6.25 µM to 35 nM (final concentration). The mixture was incubated for 1 h at room temperature. Next, the anti-6xHis beads were added to the mixture in the dark and the mixture was incubated for another hour. Finally, the streptavidin beads were added, followed by another hour of incubation. All the incubation steps were performed with the plate sealed and protected from the light. The final volume of each well was 20 µl. The plate was then read in a PHERAstar FS (BMG LABTECH). Each of the competitors was titrated in quadruplicate. The fitting and IC_{50} determination were performed in GraphPad Prism 7 (GraphPad Software, USA).

CHAPTER 7

REFERENCES

1. Edfeldt, F. N. B., Folmer, R. H. A. & Breeze, A. L. Fragment screening to predict druggability (ligandability) and lead discovery success. *Drug Discov. Today* **16,** 284–287 (2011).

2. Hopkins, A. L. & Groom, C. R. The druggable genome. *Nat Rev Drug Discov* **1,** 727–730 (2002).

3. Bonetta, L. Protein-protein interactions: Interactome under construction. *Nature* **468,** 851–854 (2010).

4. Surade, S. & Blundell, T. L. Structural biology and drug discovery of difficult targets: the limits of ligandability. *Chem. Biol.* **19,** 42–50 (2012).

5. Zanzoni, A., Soler-López, M. & Aloy, P. A network medicine approach to human disease. *FEBS Lett.* **583,** 1759–1765 (2009).

6. Perrakis, A., Musacchio, A., Cusack, S. & Petosa, C. Investigating a macromolecular complex: the toolkit of methods. *J. Struct. Biol.* **175,** 106–112 (2011).

7. Cooper, G. M. *The Cell: a molecular approach*. (Sinauer Associates, Inc., 2000).

8. Amm, I., Sommer, T. & Wolf, D. H. Protein quality control and elimination of protein waste: the role of the ubiquitin-proteasome system. *Biochim. Biophys. Acta* **1843,** 182–196 (2014).

9. *The Nobel Prize in Chemistry 2004.* (Nobelprize.org.). at <http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2004/>

10. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **67,** 425–479 (1998).

11. Schulman, B. A. & Harper, J. W. Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat. Rev. Mol. Cell Biol.* **10,** 319–331 (2009).

12. Deshaies, R. J. & Joazeiro, C. A. P. RING domain E3 ubiquitin ligases. *Annu. Rev. Biochem.* **78,** 399–434 (2009).

13. Komander, D. & Rape, M. The ubiquitin code. *Annu. Rev. Biochem.* **81,** 203–229 (2012).

14. Vittal, V., Stewart, M. D., Brzovic, P. S. & Klevit, R. E. Regulating the Regulators: Recent Revelations in the Control of E3 Ubiquitin Ligases. *J. Biol. Chem.* **290,** 21244–21251 (2015).

15. Buetow, L. & Huang, D. T. Structural insights into the catalysis and regulation of E3 ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **17,** 626–642 (2016).

16. Petroski, M. D. & Deshaies, R. J. Function and regulation of cullin-RING ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **6,** 9–20 (2005).

17. Petroski, M. D. & Deshaies, R. J. Function and regulation of cullin-RING ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **6,** 9–20 (2005).

18. Duda, D. M., Scott, D. C., Calabrese, M. F., Zimmerman, E. S., Zheng, N. & Schulman, B. A. Structural regulation of cullin-RING ubiquitin ligase complexes. *Current Opinion in Structural Biology* **21,** 257–264 (2011).

19. Duda, D. M., Borg, L. A., Scott, D. C., Hunt, H. W., Hammel, M. & Schulman, B. A. Structural insights into NEDD8 activation of cullin-RING ligases: conformational control of conjugation. *Cell* **134,** 995–1006 (2008).

20. Lingaraju, G. M., Bunker, R. D., Cavadini, S., Hess, D., Hassiepen, U., Renatus, M., Fischer, E. S. & Thomä, N. H. Crystal structure of the human COP9 signalosome. *Nature* **512,** 161–165 (2014).

21. Morimoto, M., Nishida, T., Nagayama, Y. & Yasuda, H. Nedd8-modification of Cul1 is promoted by Roc1 as a Nedd8-E3 ligase and

regulates its stability. *Biochem. Biophys. Res. Commun.* **301,** 392–398 (2003).

22. Mosadeghi, R., Reichermeier, K. M., Winkler, M., Schreiber, A., Reitsma, J. M., Zhang, Y., Stengel, F., Cao, J., Kim, M., Sweredoski, M. J., Hess, S., Leitner, A., Aebersold, R., Peter, M., Deshaies, R. J. & Enchev, R. I. Structural and kinetic analysis of the COP9-Signalosome activation and the cullin-RING ubiquitin ligase deneddylation cycle. *Elife* **5,** E2 (2016).

23. Read, M. A., Brownell, J. E., Gladysheva, T. B., Hottelet, M., Parent, L. A., Coggins, M. B., Pierce, J. W., Podust, V. N., Luo, R. S., Chau, V. & Palombella, V. J. Nedd8 modification of cul-1 activates SCF(beta(TrCP))-dependent ubiquitination of IkappaBalpha. *Mol. Cell. Biol.* **20,** 2326–2333 (2000).

24. Chen, Z., Sui, J., Zhang, F. & Zhang, C. Cullin family proteins and tumorigenesis: genetic association and molecular mechanisms. *J Cancer* **6,** 233–242 (2015).

25. Bulatov, E. & Ciulli, A. Targeting Cullin-RING E3 ubiquitin ligases for drug discovery: structure, assembly and small-molecule modulation. *Biochem. J.* **467,** 365–386 (2015).

26. Latif, F., Tory, K., Gnarra, J., Yao, M., Duh, F. M., Orcutt, M. L., Stackhouse, T., Kuzmin, I., Modi, W. & Geil, L. Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science* **260,** 1317–1320 (1993).

27. Ohh, M., Yauch, R. L., Lonergan, K. M., Whaley, J. M., Stemmer-Rachamimov, A. O., Louis, D. N., Gavin, B. J., Kley, N., Kaelin, W. G. & Iliopoulos, O. The von Hippel-Lindau tumor suppressor protein is

required for proper assembly of an extracellular fibronectin matrix. *Mol. Cell* **1,** 959–968 (1998).

28. Okumura, F., Matsuzaki, M., Nakatsukasa, K. & Kamura, T. The Role of Elongin BC-Containing Ubiquitin Ligases. *Front Oncol* **2,** 10 (2012).

29. Conaway, J. W., Kamura, T. & Conaway, R. C. The Elongin BC complex and the von Hippel-Lindau tumor suppressor protein. *Biochim. Biophys. Acta* **1377,** M49–54 (1998).

30. Zhou, W., Wei, W. & Sun, Y. Genetically engineered mouse models for functional studies of SKP1-CUL1-F-box-protein (SCF) E3 ubiquitin ligases. *Cell Res.* **23,** 599–619 (2013).

31. Wei, D. & Sun, Y. Small RING Finger Proteins RBX1 and RBX2 of SCF E3 Ubiquitin Ligases: The Role in Cancer and as Cancer Targets. *Genes Cancer* **1,** 700–707 (2010).

32. Kamura, T., Conrad, M. N., Yan, Q., Conaway, R. C. & Conaway, J. W. The Rbx1 subunit of SCF and VHL E3 ubiquitin ligase activates Rub1 modification of cullins Cdc53 and Cul2. *Genes Dev.* **13,** 2928–2933 (1999).

33. Cai, W. & Yang, H. The structure and regulation of Cullin 2 based E3 ubiquitin ligases and their biological functions. *Cell Div* **11,** 7 (2016).

34. Anderson, K., Nordquist, K. A., Gao, X., Hicks, K. C., Zhai, B., Gygi, S. P. & Patel, T. B. Regulation of cellular levels of Sprouty2 protein by prolyl hydroxylase domain and von Hippel-Lindau proteins. *J. Biol. Chem.* **286,** 42027–42036 (2011).

35. Kuznetsova, A. V., Meller, J., Schnell, P. O., Nash, J. A., Ignacak, M. L., Sanchez, Y., Conaway, J. W., Conaway, R. C. & Czyzyk-Krzeska, M. F. von Hippel-Lindau protein binds hyperphosphorylated large subunit of

RNA polymerase II through a proline hydroxylation motif and targets it for ubiquitination. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 2706–2711 (2003).

36. Mikhaylova, O., Ignacak, M. L., Barankiewicz, T. J., Harbaugh, S. V., Yi, Y., Maxwell, P. H., Schneider, M., Van Geyte, K., Carmeliet, P., Revelo, M. P., Wyder, M., Greis, K. D., Meller, J. & Czyzyk-Krzeska, M. F. The von Hippel-Lindau tumor suppressor protein and Egl-9-Type proline hydroxylases regulate the large subunit of RNA polymerase II in response to oxidative stress. *Mol. Cell. Biol.* **28,** 2701–2717 (2008).

37. Na, X., Duan, H. O., Messing, E. M., Schoen, S. R., Ryan, C. K., di Sant'Agnese, P. A., Golemis, E. A. & Wu, G. Identification of the RNA polymerase II subunit hsRPB7 as a novel target of the von Hippel-Lindau protein. *EMBO J.* **22,** 4249–4259 (2003).

38. Zhou, L. & Yang, H. The von Hippel-Lindau tumor suppressor protein promotes c-Cbl-independent poly-ubiquitylation and degradation of the activated EGFR. *PLoS ONE* **6,** e23936 (2011).

39. Chan, D. A., Sutphin, P. D., Yen, S.-E. & Giaccia, A. J. Coordinate regulation of the oxygen-dependent degradation domains of hypoxia-inducible factor 1 alpha. *Mol. Cell. Biol.* **25,** 6415–6426 (2005).

40. Maxwell, P. H., Wiesener, M. S., Chang, G. W., Clifford, S. C., Vaux, E. C., Cockman, M. E., Wykoff, C. C., Pugh, C. W., Maher, E. R. & Ratcliffe, P. J. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* **399,** 271–275 (1999).

41. Semenza, G. L. Life with oxygen. *Science* **318,** 62–64 (2007).

42. Muchnik, E. & Kaplan, J. HIF prolyl hydroxylase inhibitors for anemia. *Expert Opin Investig Drugs* **20,** 645–656 (2011).

43. Galdeano, C., Gadd, M. S., Soares, P., Scaffidi, S., Van Molle, I., Birced, I., Hewitt, S., Dias, D. M. & Ciulli, A. Structure-guided design and optimization of small molecules targeting the protein-protein interaction between the von Hippel-Lindau (VHL) E3 ubiquitin ligase and the hypoxia inducible factor (HIF) alpha subunit with in vitro nanomolar affinities. *J. Med. Chem.* **57,** 8657–8663 (2014).

44. Frost, J., Galdeano, C., Soares, P., Gadd, M. S., Grzes, K. M., Ellis, L., Epemolu, O., Shimamura, S., Bantscheff, M., Grandi, P., Read, K. D., Cantrell, D. A., Rocha, S. & Ciulli, A. Potent and selective chemical probe of hypoxic signalling downstream of HIF-α hydroxylation via VHL inhibition. *Nat Commun* **7,** 13312 (2016).

45. Zengerle, M., Chan, K.-H. & Ciulli, A. Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4. *ACS Chem. Biol.* **10,** 1770–1777 (2015).

46. Buckley, D. L., Raina, K., Darricarrere, N., Hines, J., Gustafson, J. L., Smith, I. E., Miah, A. H., Harling, J. D. & Crews, C. M. HaloPROTACS: Use of Small Molecule PROTACs to Induce Degradation of HaloTag Fusion Proteins. *ACS Chem. Biol.* **10,** 1831–1837 (2015).

47. Fulcher, L. J., Macartney, T., Bozatzi, P., Hornberger, A., Rojas-Fernandez, A. & Sapkota, G. P. An affinity-directed protein missile system for targeted proteolysis. *Open Biol* **6,** 160255 (2016).

48. Lai, A. C. & Crews, C. M. Induced protein degradation: an emerging drug discovery paradigm. *Nat Rev Drug Discov* **16,** 101–114 (2017).

49. Perkins, J. R., Diboun, I., Dessailly, B. H., Lees, J. G. & Orengo, C. Transient protein-protein interactions: structural, functional, and network properties. *Structure* **18,** 1233–1243 (2010).

50. Milroy, L.-G., Grossmann, T. N., Hennig, S., Brunsveld, L. & Ottmann, C. Modulators of protein-protein interactions. *Chem. Rev.* **114,** 4695–4748 (2014).

51. Arkin, M. R., Tang, Y. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing toward the reality. *Chem. Biol.* **21,** 1102–1114 (2014).

52. Van Molle, I., Thomann, A., Buckley, D. L., So, E. C., Lang, S., Crews, C. M. & Ciulli, A. Dissecting fragment-based lead discovery at the von Hippel-Lindau protein:hypoxia inducible factor 1α protein-protein interface. *Chem. Biol.* **19,** 1300–1312 (2012).

53. Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J. & Pavletich, N. P. Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **274,** 948–953 (1996).

54. Rickert, M., Wang, X., Boulanger, M. J., Goriatcheva, N. & Garcia, K. C. The structure of interleukin-2 complexed with its alpha receptor. *Science* **308,** 1477–1480 (2005).

55. Arkin, M. R., Randal, M., DeLano, W. L., Hyde, J., Luong, T. N., Oslob, J. D., Raphael, D. R., Taylor, L., Wang, J., McDowell, R. S., Wells, J. A. & Braisted, A. C. Binding of small molecules to an adaptive protein-protein interface. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 1603–1608 (2003).

56. Bowman, A. L., Nikolovska-Coleska, Z., Zhong, H., Wang, S. & Carlson, H. A. Small molecule inhibitors of the MDM2-p53 interaction discovered by ensemble-based receptor models. *J. Am. Chem. Soc.* **129,** 12809–12814 (2007).

57. Vassilev, L. T., Vu, B. T., Graves, B., Carvajal, D., Podlaski, F., Filipovic,

Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C., Fotouhi, N. & Liu, E. A. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303,** 844–848 (2004).

58. Buckley, D. L., Van Molle, I., Gareiss, P. C., Tae, H. S., Michel, J., Noblin, D. J., Jorgensen, W. L., Ciulli, A. & Crews, C. M. Targeting the von Hippel-Lindau E3 ubiquitin ligase using small molecules to disrupt the VHL/HIF-1α interaction. *J. Am. Chem. Soc.* **134,** 4465–4468 (2012).

59. Nevola, L. & Giralt, E. Modulating protein-protein interactions: the potential of peptides. *Chem. Commun. (Camb.)* **51,** 3302–3315 (2015).

60. Vlieghe, P., Lisowski, V., Martinez, J. & Khrestchatisky, M. Synthetic therapeutic peptides: science and market. *Drug Discov. Today* **15,** 40–56 (2010).

61. Cromm, P. M., Spiegel, J. & Grossmann, T. N. Hydrocarbon stapled peptides as modulators of biological function. *ACS Chem. Biol.* **10,** 1362–1375 (2015).

62. Azzarito, V., Long, K., Murphy, N. S. & Wilson, A. J. Inhibition of α-helix-mediated protein–protein interactions using designed molecules. *Nature Chem* **5,** 161–173 (2013).

63. Zorzi, A., Deyle, K. & Heinis, C. Cyclic peptide therapeutics: past, present and future. *Curr Opin Chem Biol* **38,** 24–29 (2017).

64. Li, P. & Roller, P. P. Cyclization strategies in peptide derived drug design. *Curr Top Med Chem* **2,** 325–341 (2002).

65. Cardote, T. A. F. & Ciulli, A. Cyclic and Macrocyclic Peptides as Chemical Tools To Recognise Protein Surfaces and Probe Protein-Protein Interactions. *ChemMedChem* **11,** 787–794 (2016).

66. Rutherford, T., Heinis, C., Freund, S. & Winter, G. Phage-encoded

combinatorial chemical libraries based on bicyclic peptides. *Nat Chem Biol* **5,** 502–507 (2009).

67. Morales-Sanfrutos, J., Angelini, A., Chen, S., Heinis, C. & Cutting, B. Structurally Diverse Cyclisation Linkers Impose Different Backbone Conformations in Bicyclic Peptides. *ChemBioChem* **13,** 1032–1038 (2012).

68. Zhou, M., Li, Q. & Wang, R. Current Experimental Methods for Characterizing Protein-Protein Interactions. *ChemMedChem* **11,** 738–756 (2016).

69. Renaud, J.-P., Chung, C.-W., Danielson, U. H., Egner, U., Hennig, M., Hubbard, R. E. & Nar, H. Biophysics in drug discovery: impact, challenges and opportunities. *Nat Rev Drug Discov* **15,** 679–698 (2016).

70. Kleywegt, G. J. & Jones, T. A. Model building and refinement practice. *Meth. Enzymol.* **277,** 208–230 (1997).

71. Dyda, F. Developments in low-resolution biological X-ray crystallography. *F1000 Biol Rep* **2,** 80 (2010).

72. Karmali, A. M., Blundell, T. L. & Furnham, N. Model-building strategies for low-resolution X-ray crystallographic data. *Acta Crystallogr. D Biol. Crystallogr.* **65,** 121–127 (2009).

73. Sibanda, B. L., Chirgadze, D. Y., Ascher, D. B. & Blundell, T. L. DNA-PKcs structure suggests an allosteric mechanism modulating DNA double-strand break repair. *Science* **355,** 520–524 (2017).

74. Misasi, J., Gilman, M. S. A., Kanekiyo, M., Gui, M., Cagigi, A., Mulangu, S., Corti, D., Ledgerwood, J. E., Lanzavecchia, A., Cunningham, J., Muyembe-Tamfun, J. J., Baxa, U., Graham, B. S., Xiang, Y., Sullivan, N. J. & McLellan, J. S. Structural and molecular basis for Ebola virus

neutralization by protective human antibodies. *Science* **351,** 1343–1346 (2016).

75. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc* **2,** 2212–2221 (2007).

76. Linkuvienė, V., Krainer, G., Chen, W.-Y. & Matulis, D. Isothermal titration calorimetry for drug design: Precision of the enthalpy and binding constant measurements and comparison of the instruments. *Anal. Biochem.* **515,** 61–64 (2016).

77. Dias, D. M. & Ciulli, A. NMR approaches in structure-based lead discovery: Recent developments and new frontiers for targeting multi-protein complexes. *Prog. Biophys. Mol. Biol.* (2014). doi:10.1016/j.pbiomolbio.2014.08.012

78. Carr, H. I. & Purcell, E. M. Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments. *Phisical Reviews* **94,**

79. Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc* **73,** 1–16 (2013).

80. Shah, N. B. & Duncan, T. M. Bio-layer interferometry for measuring kinetics of protein-protein interactions and allosteric ligand effects. *J Vis Exp* e51383 (2014). doi:10.3791/51383

81. Bielefeld-Sevigny, M. AlphaLISA immunoassay platform- the 'no-wash' high-throughput alternative to ELISA. *Assay Drug Dev Technol* **7,** 90–92 (2009).

82. Nguyen, H. C., Yang, H., Fribourgh, J. L., Wolfe, L. S. & Xiong, Y. Insights into Cullin-RING E3 Ubiquitin Ligase Recruitment: Structure of the VHL-EloBC-Cul2 Complex. *Structure* (2015).

doi:10.1016/j.str.2014.12.014

83. Lydeard, J. R., Schulman, B. A. & Harper, J. W. Building and remodelling Cullin-RING E3 ubiquitin ligases. *EMBO Rep.* **14,** 1050–1061 (2013).

84. Zheng, N., Schulman, B. A., Song, L., Miller, J. J., Jeffrey, P. D., Wang, P., Chu, C., Koepp, D. M., Elledge, S. J., Pagano, M., Conaway, R. C., Conaway, J. W., Harper, J. W. & Pavletich, N. P. Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416,** 703–709 (2002).

85. Fischer, E. S., Scrima, A., Böhm, K., Matsumoto, S., Lingaraju, G. M., Faty, M., Yasuda, T., Cavadini, S., Wakasugi, M., Hanaoka, F., Iwai, S., Gut, H., Sugasawa, K. & Thomä, N. H. The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell* **147,** 1024–1039 (2011).

86. Lucas, X. & Ciulli, A. Recognition of substrate degrons by E3 ubiquitin ligases and modulation by small-molecule mimicry strategies. *Current Opinion in Structural Biology* **44,** 101–110 (2017).

87. Winter, G. E., Buckley, D. L., Paulk, J., Roberts, J. M., Souza, A., Dhe-Paganon, S. & Bradner, J. E. DRUG DEVELOPMENT. Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science* **348,** 1376–1381 (2015).

88. Lu, J., Qian, Y., Altieri, M., Dong, H., Wang, J., Raina, K., Hines, J., Winkler, J. D., Crew, A. P., Coleman, K. & Crews, C. M. Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. *Chem. Biol.* **22,** 755–763 (2015).

89. Lee, D. W., Peggie, M., Deak, M., Toth, R., Gage, Z. O., Wood, N.,

Schilde, C., Kurz, T. & Knebel, A. The Dac-tag, an affinity tag based on penicillin-binding protein 5. *Anal. Biochem.* **428,** 64–72 (2012).

90. Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol.* **33,** 491–497 (1968).

91. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25,** 1605–1612 (2004).

92. Goldenberg, S. J., Cascio, T. C., Shumway, S. D., Garbutt, K. C., Liu, J., Xiong, Y. & Zheng, N. Structure of the Cand1-Cul1-Roc1 complex reveals regulatory mechanisms for the assembly of the multisubunit cullin-dependent ubiquitin ligases. *Cell* **119,** 517–528 (2004).

93. Kim, Y. K., Kwak, M.-J., Ku, B., Suh, H.-Y., Joo, K., Lee, J., Jung, J. U. & Oh, B.-H. Structural basis of intersubunit recognition in elongin BC-cullin 5-SOCS box ubiquitin-protein ligase complexes. *Acta Crystallogr. D Biol. Crystallogr.* **69,** 1587–1597 (2013).

94. Keegan, R. M. & Winn, M. D. Automated search-model discovery and preparation for structure solution by molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* **63,** 447–457 (2007).

95. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. Phaser crystallographic software. *J Appl Crystallogr* **40,** 658–674 (2007).

96. Spratt, D. E., Wu, K., Kovacev, J., Pan, Z.-Q. & Shaw, G. S. Selective recruitment of an E2~ubiquitin complex by an E3 ubiquitin ligase. *J. Biol. Chem.* **287,** 17374–17385 (2012).

97. Kamura, T., Maenaka, K., Kotoshiba, S., Matsumoto, M., Kohda, D.,

Conaway, R. C., Conaway, J. W. & Nakayama, K. I. VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. *Genes Dev.* **18,** 3055–3065 (2004).

98. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F. & Murshudov, G. N. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2184–2195 (2004).

99. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 213–221 (2010).

100. Nicholls, R. A., Long, F. & Murshudov, G. N. Low-resolution refinement tools in REFMAC5. *Acta Crystallogr. D Biol. Crystallogr.* **68,** 404–417 (2012).

101. Liu, J. & Nussinov, R. Flexible cullins in cullin-RING E3 ligases allosterically regulate ubiquitination. *J. Biol. Chem.* **286,** 40934–40942 (2011).

102. Canning, P., Cooper, C. D. O., Krojer, T., Murray, J. W., Pike, A. C. W., Chaikuad, A., Keates, T., Thangaratnarajah, C., Hojzan, V., Ayinampudi, V., Marsden, B. D., Gileadi, O., Knapp, S., Delft, von, F. & Bullock, A. N. Structural basis for Cul3 protein assembly with the BTB-Kelch family of E3 ubiquitin ligases. *J. Biol. Chem.* **288,** 7803–7814 (2013).

103. Angers, S., Li, T., Yi, X., MacCoss, M. J., Moon, R. T. & Zheng, N. Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. *Nature* **443,** 590–593 (2006).

104. Calabrese, M. F., Scott, D. C., Duda, D. M., Grace, C. R. R., Kurinov, I., Kriwacki, R. W. & Schulman, B. A. A RING E3-substrate complex poised for ubiquitin-like protein transfer: structural insights into cullin-RING ligases. *Nat. Struct. Mol. Biol.* **18,** 947–949 (2011).

105. Onel, M., Sumbul, F., Liu, J., Nussinov, R. & Haliloglu, T. Cullin Neddylation May Allosterically Tune Polyubiquitin Chain Length and Topology. *Biochem. J.* BCJ20160748 (2017). doi:10.1042/BCJ20160748

106. Scott, D. C., Sviderskiy, V. O., Monda, J. K., Lydeard, J. R., Cho, S. E., Harper, J. W. & Schulman, B. A. Structure of a RING E3 Trapped in Action Reveals Ligation Mechanism for the Ubiquitin-like Protein NEDD8. *Cell* **157,** 1671–1684 (2014).

107. Duda, D. M., Olszewski, J. L., Tron, A. E., Hammel, M., Lambert, L. J., Waddell, M. B., Mittag, T., DeCaprio, J. A. & Schulman, B. A. Structure of a glomulin-RBX1-CUL1 complex: inhibition of a RING E3 ligase through masking of its E2-binding surface. *Mol. Cell* **47,** 371–382 (2012).

108. Morimoto, M., Nishida, T., Honda, R. & Yasuda, H. Modification of cullin-1 by ubiquitin-like protein Nedd8 enhances the activity of SCF(skp2) toward p27(kip1). *Biochem. Biophys. Res. Commun.* **270,** 1093–1096 (2000).

109. Podust, V. N., Brownell, J. E., Gladysheva, T. B., Luo, R. S., Wang, C., Coggins, M. B., Pierce, J. W., Lightcap, E. S. & Chau, V. A Nedd8

conjugation pathway is essential for proteolytic targeting of p27Kip1 by ubiquitination. *Proc. Natl. Acad. Sci. U.S.A.* **97,** 4579–4584 (2000).

110. Wu, K., Chen, A. & Pan, Z. Q. Conjugation of Nedd8 to CUL1 enhances the ability of the ROC1-CUL1 complex to promote ubiquitin polymerization. *J. Biol. Chem.* **275,** 32317–32324 (2000).

111. Kurz, T., Chou, Y.-C., Willems, A. R., Meyer-Schaller, N., Hecht, M.-L., Tyers, M., Peter, M. & Sicheri, F. Dcn1 functions as a scaffold-type E3 ligase for cullin neddylation. *Mol. Cell* **29,** 23–35 (2008).

112. Chiba, T. In vitro systems for NEDD8 conjugation by Ubc12. *Meth. Enzymol.* **398,** 68–73 (2005).

113. Morimoto, M., Nishida, T., Nagayama, Y. & Yasuda, H. Nedd8-modification of Cul1 is promoted by Roc1 as a Nedd8-E3 ligase and regulates its stability. *Biochem. Biophys. Res. Commun.* **301,** 392–398 (2003).

114. Muniz, J. R. C., Guo, K., Kershaw, N. J., Ayinampudi, V., Delft, von, F., Babon, J. J. & Bullock, A. N. Molecular architecture of the ankyrin SOCS box family of Cul5-dependent E3 ubiquitin ligases. *J. Mol. Biol.* **425,** 3166–3177 (2013).

115. Thomas, J. C., Matak-Vinkovic, D., Van Molle, I. & Ciulli, A. Multimeric complexes among ankyrin-repeat and SOCS-box protein 9 (ASB9), ElonginBC, and Cullin 5: insights into the structure and assembly of ECS-type Cullin-RING E3 ubiquitin ligases. *Biochemistry* **52,** 5236–5246 (2013).

116. Bulatov, E., Martin, E. M., Chatterjee, S., Knebel, A., Shimamura, S., Konijnenberg, A., Johnson, C., Zinn, N., Grandi, P., Sobott, F. & Ciulli, A. Biophysical studies on interactions and assembly of full-size E3

ubiquitin ligase: suppressor of cytokine signaling 2 (SOCS2)-elongin BC-cullin 5-ring box protein 2 (RBX2). *J. Biol. Chem.* **290,** 4178–4191 (2015).

117. Babon, J. J., Sabo, J. K., Zhang, J.-G., Nicola, N. A. & Norton, R. S. The SOCS box encodes a hierarchy of affinities for Cullin5: implications for ubiquitin ligase formation and cytokine signalling suppression. *J. Mol. Biol.* **387,** 162–174 (2009).

118. Salter, J. D., Lippa, G. M., Belashov, I. A. & Wedekind, J. E. Core-binding factor β increases the affinity between human Cullin 5 and HIV-1 Vif within an E3 ligase complex. *Biochemistry* **51,** 8702–8704 (2012).

119. Prabhu, N. V. & Sharp, K. A. Heat capacity in proteins. *Annu Rev Phys Chem* **56,** 521–548 (2005).

120. Spolar, R. S., Livingstone, J. R. & Record, M. T. Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry* **31,** 3947–3955 (1992).

121. Robertson, A. D. & Murphy, K. P. Protein Structure and the Energetics of Protein Stability. *Chem. Rev.* **97,** 1251–1268 (1997).

122. Myers, J. K., Pace, C. N. & Scholtz, J. M. Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4,** 2138–2148 (1995).

123. Makhatadze, G. I. & Privalov, P. L. Energetics of protein structure. *Adv. Protein Chem.* **47,** 307–425 (1995).

124. Robertson, A. D. & Murphy, K. P. Protein Structure and the Energetics of Protein Stability. *Chem. Rev.* **97,** 1251–1268 (1997).

125. Murphy, K. P. & Freire, E. Thermodynamics of structural stability and

cooperative folding behavior in proteins. *Adv. Protein Chem.* **43,** 313–361 (1992).

126. Yasgar, A., Jadhav, A., Simeonov, A. & Coussens, N. P. AlphaScreen-Based Assays: Ultra-High-Throughput Screening for Small-Molecule Inhibitors of Challenging Enzymes and Protein-Protein Interactions. *Methods Mol. Biol.* **1439,** 77–98 (2016).

127. Mahrour, N., Redwine, W. B., Florens, L., Swanson, S. K., Martin-Brown, S., Bradford, W. D., Staehling-Hampton, K., Washburn, M. P., Conaway, R. C. & Conaway, J. W. Characterization of Cullin-box sequences that direct recruitment of Cul2-Rbx1 and Cul5-Rbx2 modules to Elongin BC-based ubiquitin ligases. *J. Biol. Chem.* **283,** 8005–8013 (2008).

128. Bullock, A. N., Debreczeni, J. E., Edwards, A. M., Sundström, M. & Knapp, S. Crystal structure of the SOCS2-elongin C-elongin B complex defines a prototypical SOCS box ubiquitin ligase. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 7637–7642 (2006).

129. Liu, J. & Nussinov, R. Allosteric effects in the marginally stable von Hippel-Lindau tumor suppressor protein and allostery-based rescue mutant design. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 901–906 (2008).

130. Liu, J. & Nussinov, R. The mechanism of ubiquitination in the cullin-RING E3 ligase machinery: conformational control of substrate orientation. *PLoS Comput. Biol.* **5,** e1000527 (2009).

131. Benyamini, H. & Friedler, A. Using peptides to study protein-protein interactions. *Future Med Chem* **2,** 989–1003 (2010).

132. Clark, E. Refolding of recombinant proteins. *Curr. Opin. Biotechnol.* **9,** 157–163 (1998).

133. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in Escherichia coli: advances and challenges. *Front Microbiol* **5,** 172 (2014).

134. Rothman, J. E. & Schekman, R. *Molecular mechanism of protein folding in the cell. Cell* **146,** 851–854 (2011).

135. Kanaya, K., Tsai, A.-L. & Kamitani, T. Cobalt- and nickel-binding property of cullin-2. *Biochem. Biophys. Res. Commun.* **290,** 294–299 (2002).

136. Bondos, S. E. & Bicknell, A. Detection and prevention of protein aggregation before, during, and after purification. *Anal. Biochem.* **316,** 223–231 (2003).

137. Li, T., Pavletich, N. P., Schulman, B. A. & Zheng, N. High-level expression and purification of recombinant SCF ubiquitin ligases. *Meth. Enzymol.* **398,** 125–142 (2005).

138. Knauth, K., Cartwright, E., Freund, S., Bycroft, M. & Buchberger, A. VHL mutations linked to type 2C von Hippel-Lindau disease cause extensive structural perturbations in pVHL. *J. Biol. Chem.* **284,** 10514–10522 (2009).

139. White, T. R., Renzelman, C. M., Rand, A. C., Rezai, T., McEwen, C. M., Gelev, V. M., Turner, R. A., Linington, R. G., Leung, S. S. F., Kalgutkar, A. S., Bauman, J. N., Zhang, Y., Liras, S., Price, D. A., Mathiowetz, A. M., Jacobson, M. P. & Lokey, R. S. On-resin N-methylation of cyclic peptides for discovery of orally bioavailable scaffolds. *Nat Chem Biol* **7,** 810–817 (2011).

140. Chen, X., Zaro, J. L. & Shen, W.-C. Fusion protein linkers: property, design and functionality. *Adv. Drug Deliv. Rev.* **65,** 1357–1369 (2013).

141. Martínez-Alonso, M., González-Montalbán, N., García-Fruitós, E. & Villaverde, A. The Functional quality of soluble recombinant polypeptides produced in Escherichia coli is defined by a wide conformational spectrum. *Appl. Environ. Microbiol.* **74,** 7431–7433 (2008).

142. Brondyk, W. H. Selecting an appropriate method for expressing a recombinant protein. *Meth. Enzymol.* **463,** 131–147 (2009).

143. Rentero Rebollo, I., Angelini, A. & Heinis, C. Phage display libraries of differently sized bicyclic peptides. *MedChemComm* 145–150 (2012).

144. Diderich, P. & Heinis, C. Directed evolution of bicyclic peptides for therapeutic application. *Chimia (Aarau)* **67,** 910–915 (2013).

145. Chen, S., Bertoldo, D., Angelini, A., Pojer, F. & Heinis, C. Peptide ligands stabilized by small molecules. *Angew. Chem. Int. Ed. Engl.* **53,** 1602–1606 (2014).

146. Dhaliwal, B., Pang, M. O. Y., Keeble, A. H., James, L. K., Gould, H. J., McDonnell, J. M., Sutton, B. J. & Beavil, A. J. IgE binds asymmetrically to its B cell receptor CD23. *Sci Rep* **7,** 45533 (2017).

147. Baeriswyl, V. & Heinis, C. Phage selection of cyclic peptide antagonists with increased stability toward intestinal proteases. *Protein Eng. Des. Sel.* **26,** 81–89 (2013).

148. Chen, S., Rentero Rebollo, I., Buth, S. A., Morales-Sanfrutos, J., Touati, J., Leiman, P. G. & Heinis, C. Bicyclic Peptide Ligands Pulled out of Cysteine-Rich Peptide Libraries. *J. Am. Chem. Soc.* **135,** 6562–6569 (2013).

149. Heinis, C., Rutherford, T., Freund, S. & Winter, G. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat Chem*

Biol **5,** 502–507 (2009).

150. Smith, G. P. & Petrenko, V. A. Phage Display. *Chem. Rev.* **97,** 391–410 (1997).

151. Rentero Rebollo, I., Rebollo, I. R. & Heinis, C. Phage selection of bicyclic peptides. *Methods* **60,** 46–54 (2013).

152. Rentero Rebollo, I. & Heinis, C. Phage selection of bicyclic peptides. *Methods* **60,** 46–54 (2013).

153. Rentero Rebollo, I., McCallin, S., Bertoldo, D., Entenza, J. M., Moreillon, P. & Heinis, C. Development of Potent and Selective S. aureus Sortase A Inhibitors Based on Peptide Macrocycles. *ACS Med Chem Lett* **7,** 606–611 (2016).

154. Chen, S., Gfeller, D., Buth, S. A., Michielin, O., Leiman, P. G. & Heinis, C. Improving binding affinity and stability of peptide ligands by substituting glycines with D-amino acids. *ChemBioChem* **14,** 1316–1322 (2013).

155. Baeriswyl, V., Calzavarini, S., Gerschheimer, C., Diderich, P., Angelillo-Scherrer, A. & Heinis, C. Development of a selective peptide macrocycle inhibitor of coagulation factor XII toward the generation of a safe antithrombotic therapy. *J. Med. Chem.* **56,** 3742–3746 (2013).

156. Diderich, P. & Heinis, C. Phage selection of bicyclic peptides binding Her2. *Tetrahedron* **70,** 7733–7739 (2014).

157. Baeriswyl, V., Calzavarini, S., Chen, S., Zorzi, A., Bologna, L., Angelillo-Scherrer, A. & Heinis, C. A Synthetic Factor XIIa Inhibitor Blocks Selectively Intrinsic Coagulation Initiation. *ACS Chem. Biol.* **10,** 1861–1870 (2015).

158. Pollaro, L., Raghunathan, S., Morales-Sanfrutos, J., Angelini, A.,

Kontos, S. & Heinis, C. Bicyclic peptides conjugated to an albumin-binding tag diffuse efficiently into solid tumors. *Mol. Cancer Ther.* molcanther.0534.2014 (2014). doi:10.1158/1535-7163.MCT-14-0534

159. Urech-Varenne, C., Radtke, F. & Heinis, C. Phage Selection of Bicyclic Peptide Ligands of the Notch1 Receptor. *ChemMedChem* **10,** 1754–1761 (2015).

160. van de Langemheen, H., Korotkovs, V., Bijl, J., Wilson, C., Kale, S. S., Heinis, C. & Liskamp, R. M. J. Polar Hinges as Functionalized Conformational Constraints in (Bi)cyclic Peptides. *ChemBioChem* **18,** 387–395 (2017).

161. Zhao, Y. & Sun, Y. Cullin-RING Ligases as attractive anti-cancer targets. *Curr. Pharm. Des.* **19,** 3215–3225 (2013).

162. Jia, L. & Sun, Y. SCF E3 ubiquitin ligases as anticancer targets. *Curr Cancer Drug Targets* **11,** 347–356 (2011).

163. Petzold, G., Fischer, E. S. & Thomä, N. H. Structural basis of lenalidomide-induced CK1α degradation by the CRL4(CRBN) ubiquitin ligase. *Nature* **532,** 127–130 (2016).

164. Deshaies, R. J. Protein degradation: Prime time for PROTACs. *Nat Chem Biol* **11,** 634–635 (2015).

165. Lamsoul, I., Uttenweiler-Joseph, S., Moog-Lutz, C. & Lutz, P. G. Cullin 5-RING E3 ubiquitin ligases, new therapeutic targets? *Biochimie* **122,** 339–347 (2016).

166. Lu, P., Bai, X.-C., Ma, D., Xie, T., Yan, C., Sun, L., Yang, G., Zhao, Y., Zhou, R., Scheres, S. H. W. & Shi, Y. Three-dimensional structure of human γ-secretase. *Nature* **512,** 166–170 (2014).

167. Kikhney, A. G. & Svergun, D. I. A practical guide to small angle X-ray

scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **589,** 2570–2577 (2015).

168. Sambrook, J. & Russell, D. W. The inoue method for preparation and transformation of competent e. Coli: 'ultra-competent' cells. *CSH Protoc* **2006,** pdb.prot3944 (2006).

169. Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., Grosdidier, A., Hernandez, C., Ioannidis, V., Kuznetsov, D., Liechti, R., Moretti, S., Mostaguir, K., Redaschi, N., Rossier, G., Xenarios, I. & Stockinger, H. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Research* **40,** W597–603 (2012).

170. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* **69,** 1204–1214 (2013).

171. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67,** 235–242 (2011).

172. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132 (2004).

173. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21 (2010).

174. Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A.,

Llinas, M., Ulrich, E. L., Markley, J. L., Ionides, J. & Laue, E. D. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59,** 687–696 (2005).

APPENDIX I

Figure AI.1 – AlphaLISA raw data of the titration of VBC and Cul2 mutants as displacers. The experiments were performed in quadruplicate and the results are an averaged value. The error bars represent the standard deviation of each point. The fitting was performed with GraphPad Prism 7 software.

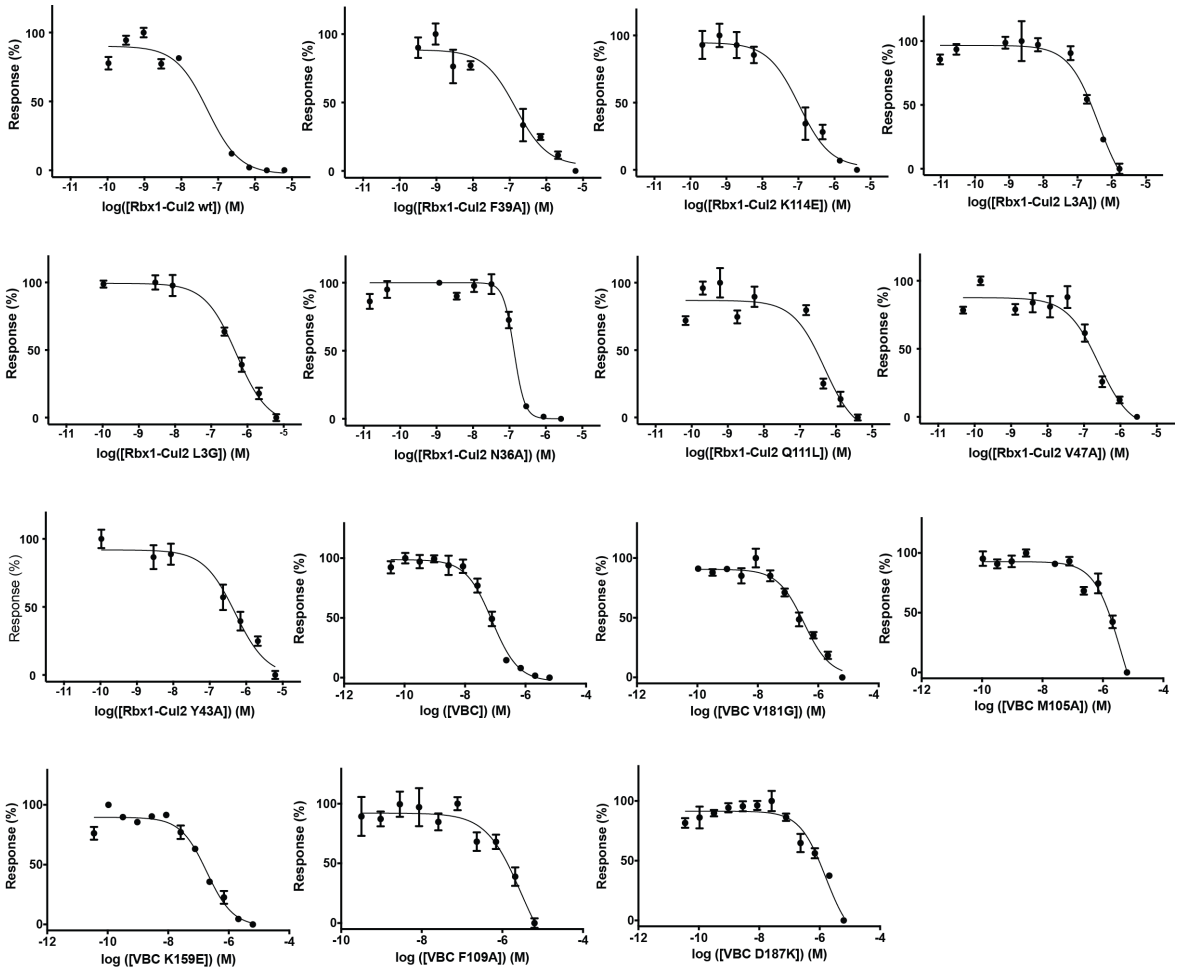


Table AI.1 – Alphalisa data of the titration of VBC and Cul2 mutants as displacers.

Displacer	IC_{50} (M)	95% CI	R^2	pIC_{50} (M)
VBC wt	7.80E-08	5.65E-08 to 1.08E-07	0.9550	-7.11 ± 0.07
VBC K159E	1.87E-07	1.34E-07 to 2.63E-07	0.9504	-6.73 ± 0.07
VBC D187K	1.59E-06	8.52E-07 to 2.97E-06	0.8773	-5.80 ± 0.13
VBC V181G	3.34E-07	2.21E-07 to 5.05E-07	0.9339	-6.48 ± 0.09
VBC M105A	3.90E-06	1.89E-06 to 8.02E-06	0.8978	-5.41 ± 0.15
VBC F109A	2.89E-06	1.40E-06 to 5.94E-06	0.8796	-5.54 ± 0.15
Cullin-2 wt	5.16E-08	2.77E-08 to 9.63E-08	0.9629	-7.29 ± 0.13
Cullin-2 L3A	4.08E-07	2.11E-07 to 7.91E-07	0.8816	-6.39 ± 0.14
Cullin-2 L3G	2.88E-07	1.26E-07 to 6.57E-07	0.8353	-6.54 ± 0.18
Cullin-2 P5A	8.11E-07	1.46E-07 to 4.50E-06	0.8302	-6.09 ± 0.36
Cullin-2 N36A	1.97E-07	1.16E-07 to 3.36E-07	0.9000	-6.71 ± 0.11
Cullin-2 F39A	1.59E-07	5.14E-08 to 4.91E-07	0.7952	-6.80 ± 0.24
Cullin-2 Y43A	2.40E-07	7.94E-08 to 7.27E-07	0.7529	-6.62 ± 0.24
Cullin-2 V47A	2.54E-07	1.43E-07 to 4.51E-07	0.8869	-6.59 ± 0.12
Cullin-2 Q111L	4.74E-07	2.27E-07 to 9.87E-07	0.8509	-6.32 ± 0.16
Cullin-2 K114E	1.04E-07	4.51E-08 to 2.41E-07	0.8798	-6.98 ± 0.18

Figure AI.2 – Isothermal Titration Calorimetry raw data of the titrations of VBC mutants *versus* Rbx1–Cul2.

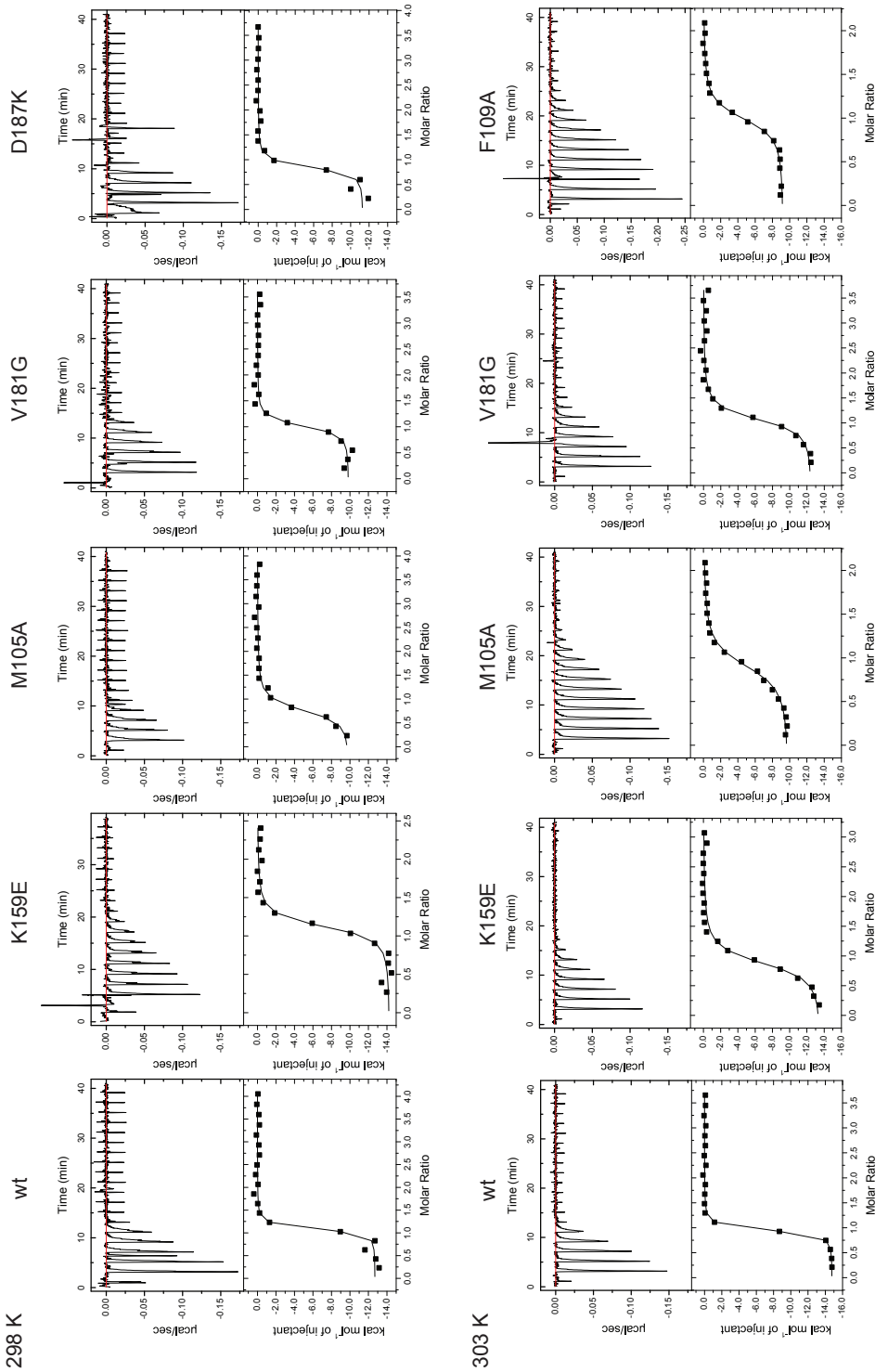


Figure AI.3 – Two-dimensional $^{1}H,^{15}N$-HSQC spectra of VBC (500 μM), 64 scans, 302K.

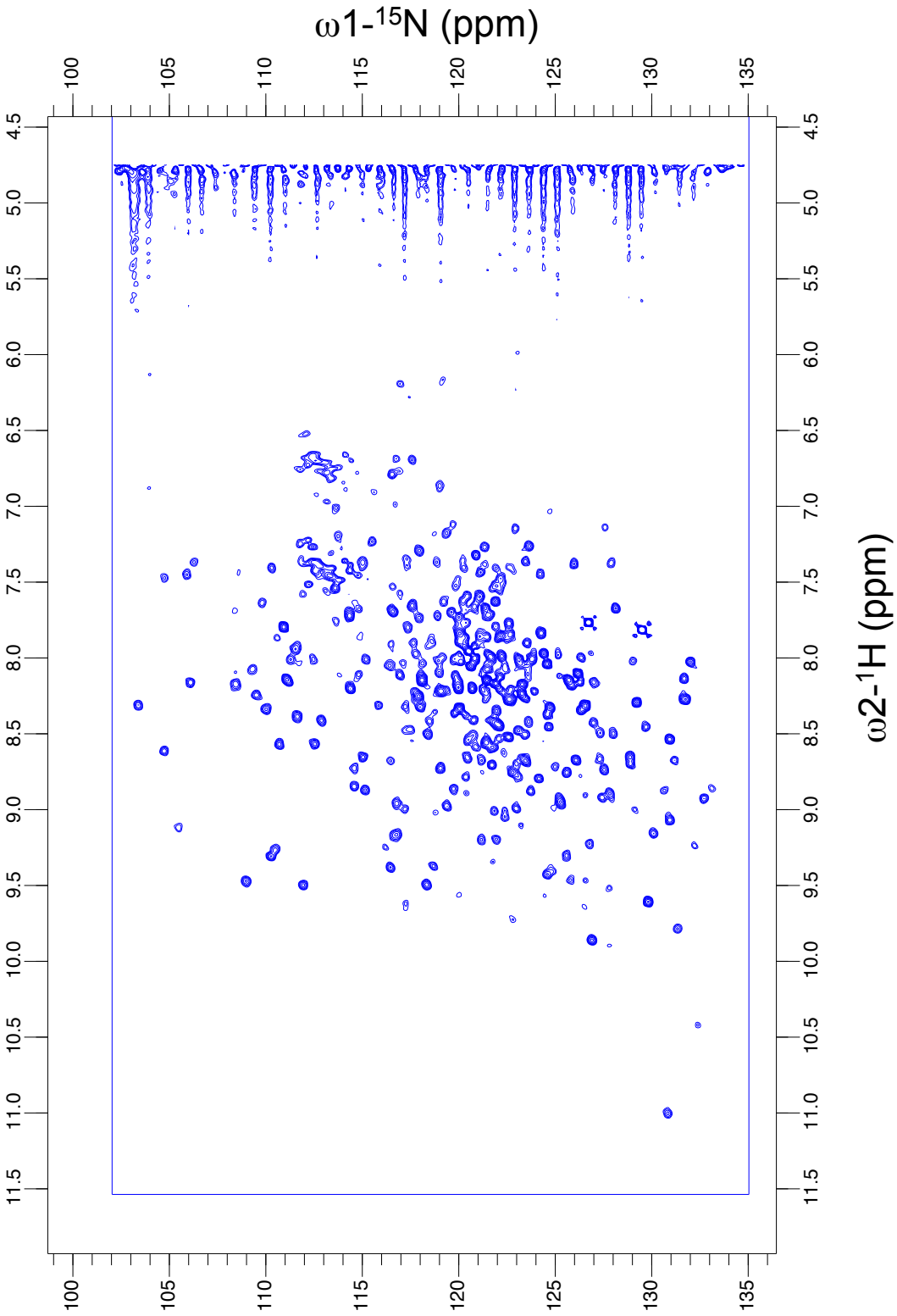


Figure AI.4 – Two-dimensional $^{1}H,^{15}N$-HSQC spectra of VBC (100 μM), 32 scans, 302 K (blue) and in the presence of 1 mM peptide **J** (grey), 2 mM peptide **J** (red) and 3 mM peptide **J** (purple).

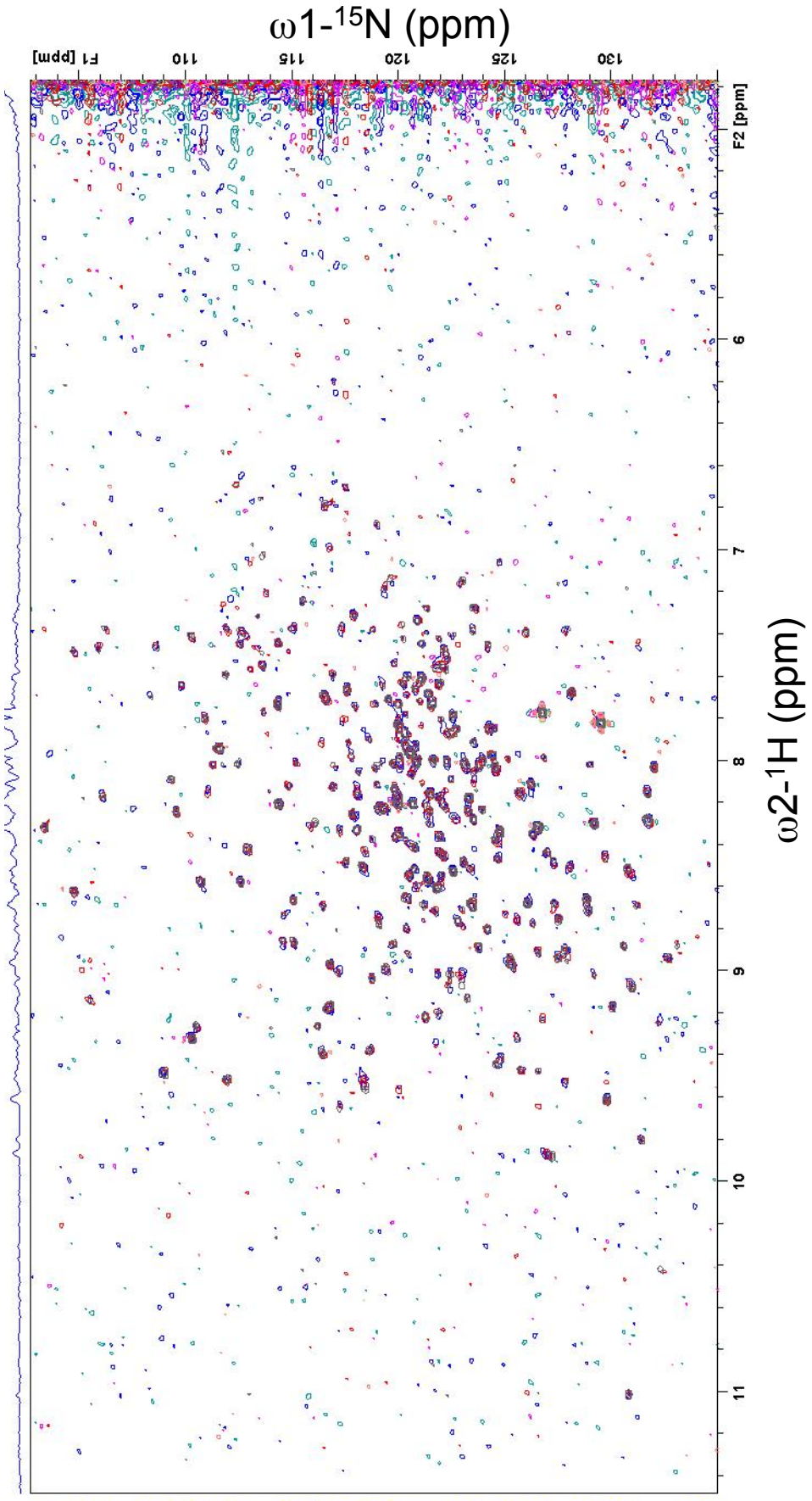


Figure AI.5 – Two-dimensional $^{1}H,^{15}N$-HSQC spectra of VBC (100 μM), 32 scans, 302 K (blue) and in the presence of 1 mM 15-mer peptide (red) and 2 mM 15-mer peptide (purple).

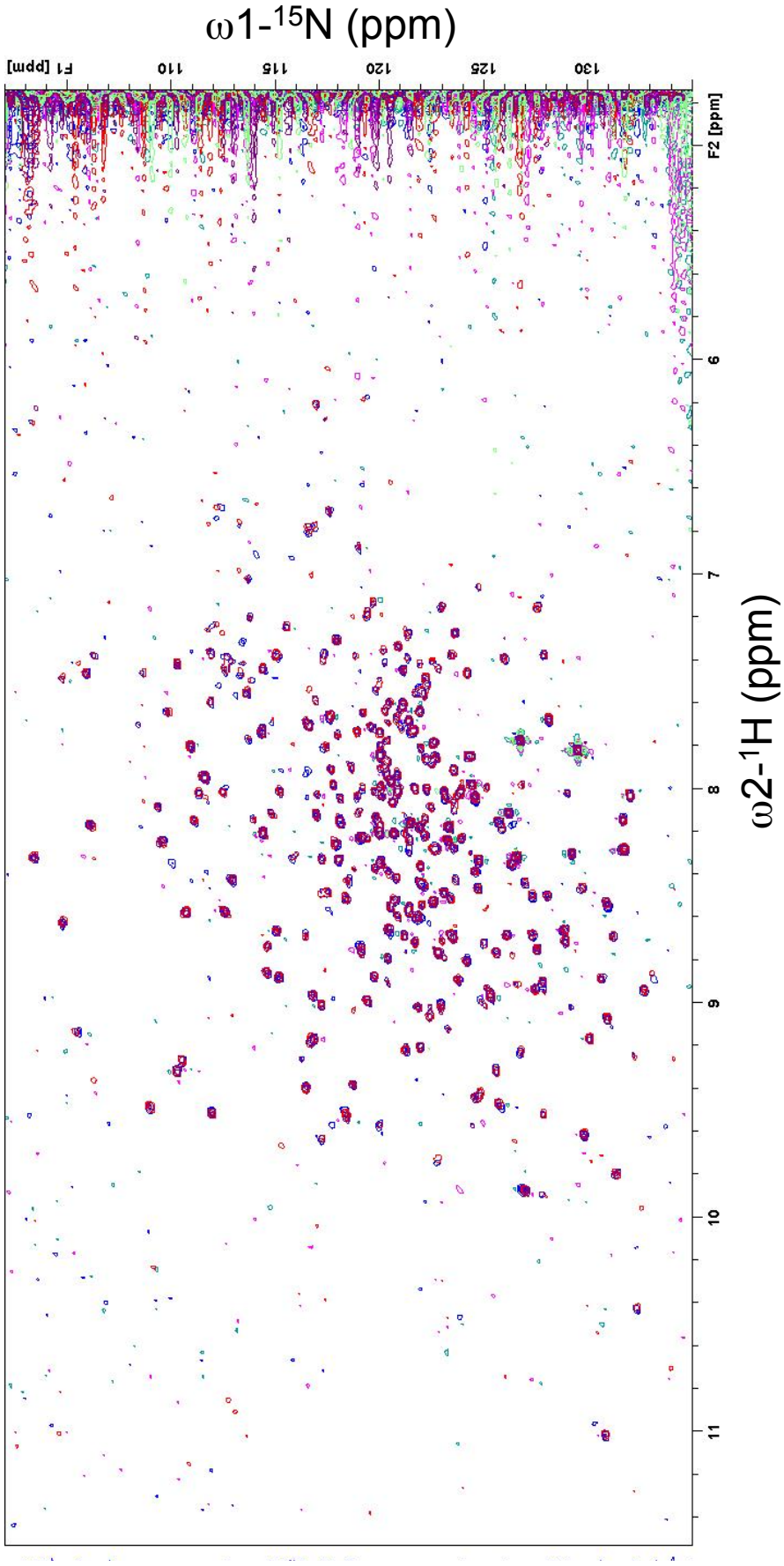
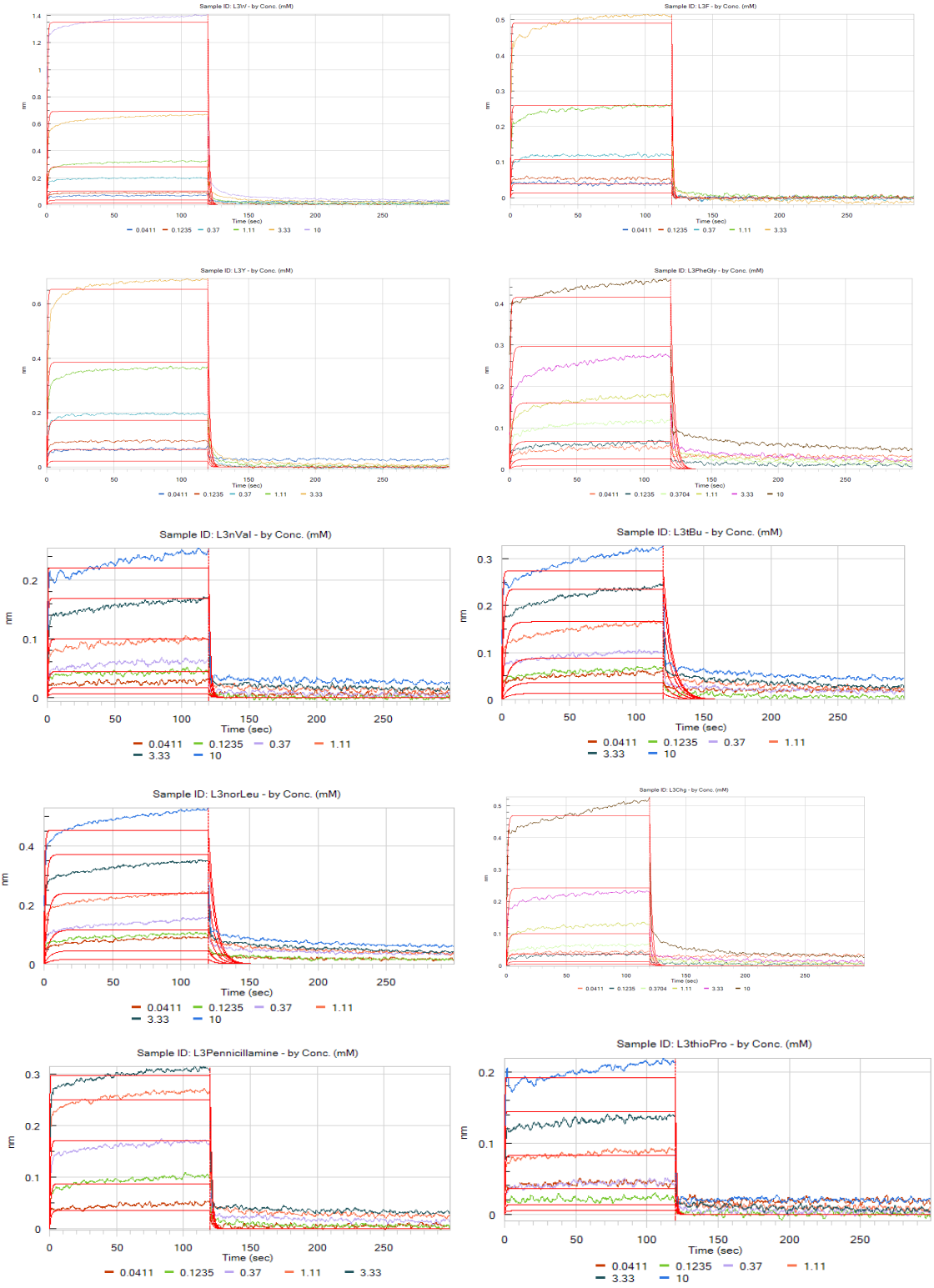
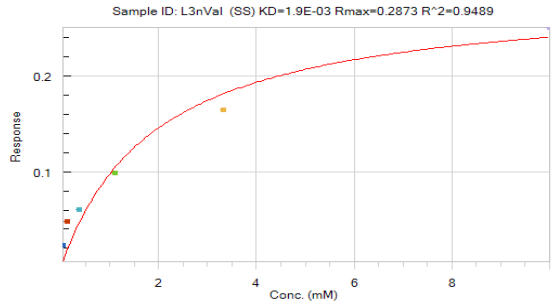
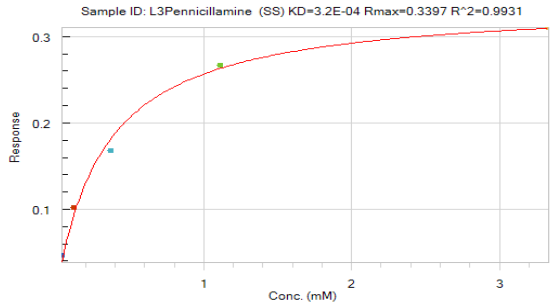
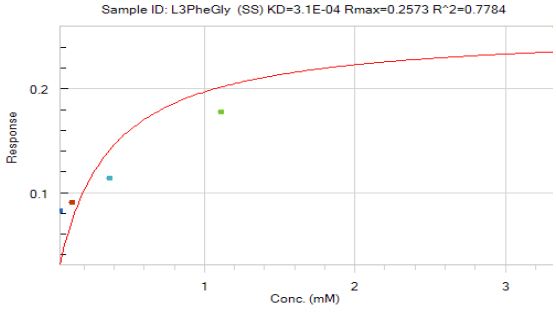


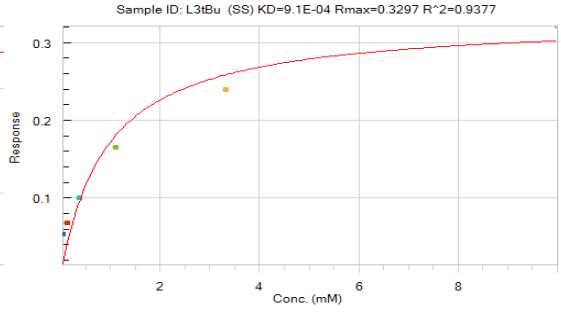
Figure AI.6 – Biolayer Interferometry data of 8-mer N-terminal Cul2 peptides with leucine replacement *versus* VBC tested on a concentration-dependent assay.

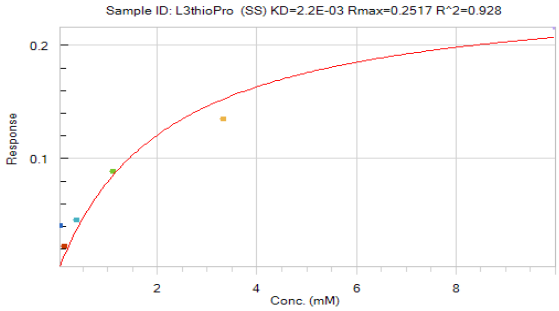


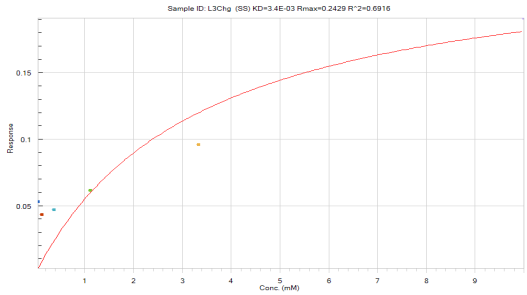

Sample ID: L3nVal (SS) KD=1.9E-03 Rmax=0.2873 R^2=0.9489
Response
Conc. (mM)

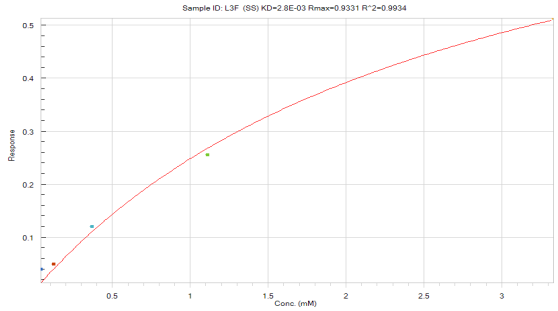

Sample ID: L3Pennicillamine (SS) KD=3.2E-04 Rmax=0.3397 R^2=0.9931
Response
Conc. (mM)

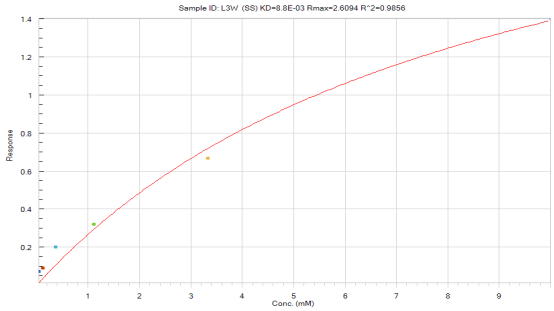

Sample ID: L3PheGly (SS) KD=3.1E-04 Rmax=0.2573 R^2=0.7784
Response
Conc. (mM)

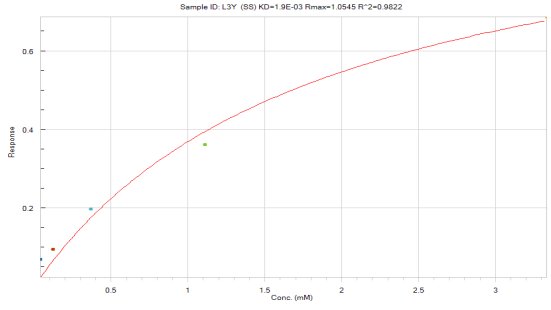

Sample ID: L3tBu (SS) KD=9.1E-04 Rmax=0.3297 R^2=0.9377
Response
Conc. (mM)


Sample ID: L3thioPro (SS) KD=2.2E-03 Rmax=0.2517 R^2=0.928
Response
Conc. (mM)


Sample ID: L3Chg (SS) KD=3.4E-03 Rmax=0.2429 R^2=0.6916
Response
Conc. (mM)


Sample ID: L3F (SS) KD=2.8E-03 Rmax=0.9331 R^2=0.9934
Response
Conc. (mM)


Sample ID: L3W (SS) KD=8.8E-03 Rmax=2.6094 R^2=0.9856
Response
Conc. (mM)


Sample ID: L3Y (SS) KD=1.9E-03 Rmax=1.0545 R^2=0.9822
Response
Conc. (mM)

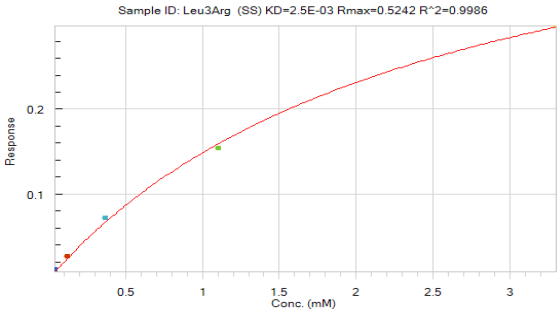
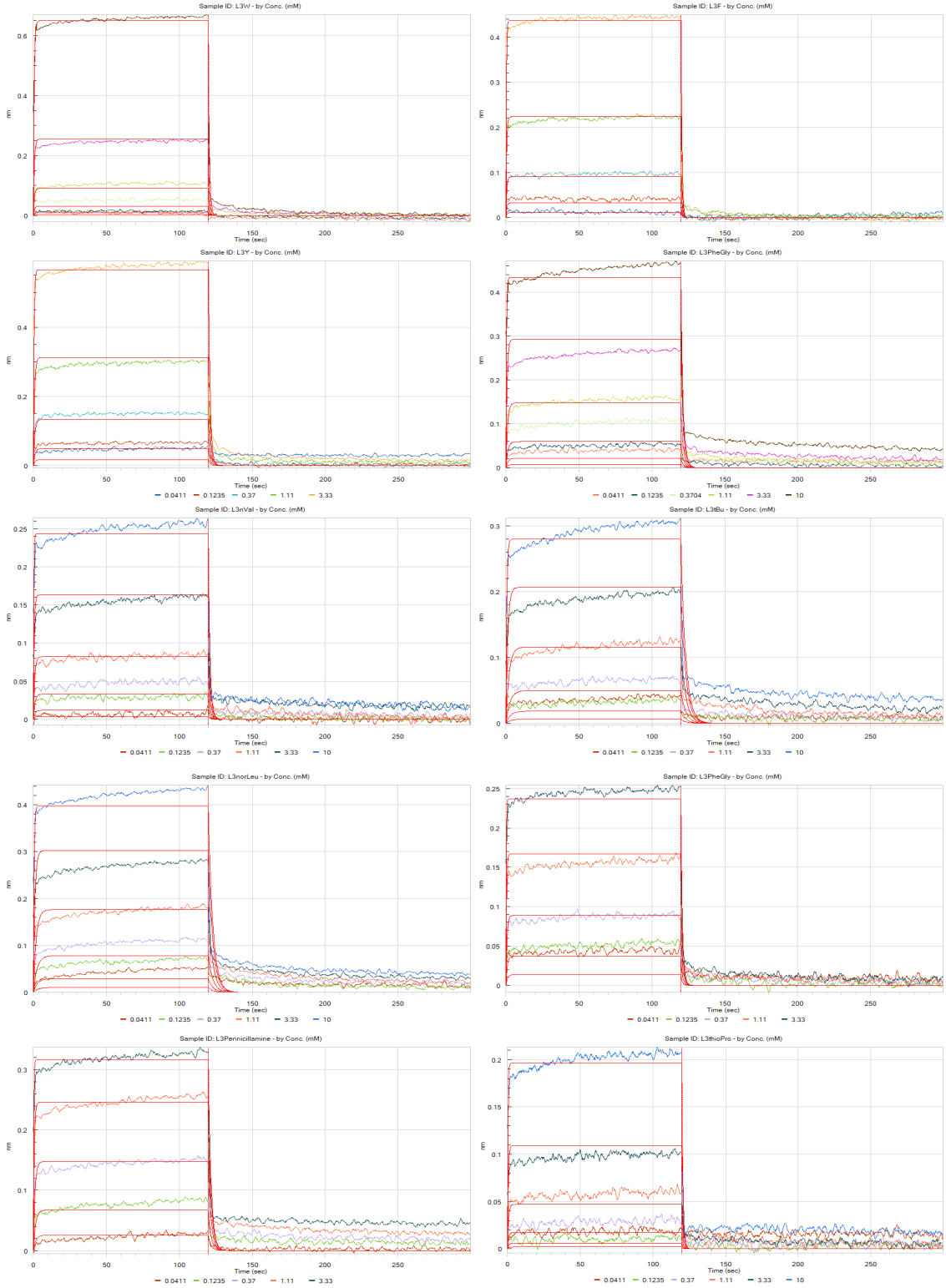
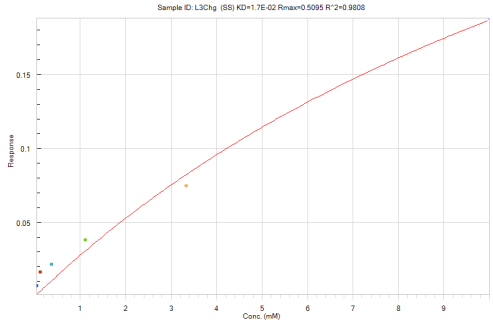
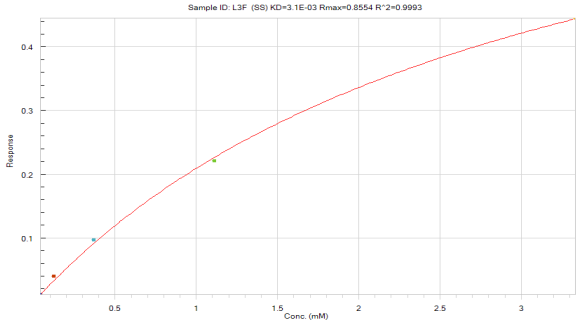
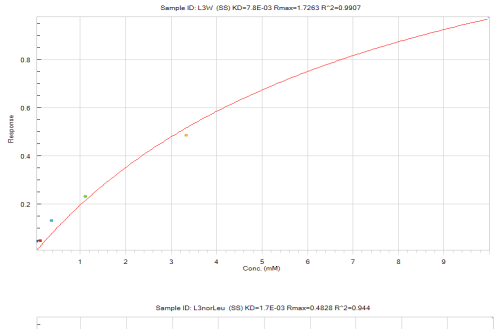

Sample ID: Leu3Arg (SS) KD=2.5E-03 Rmax=0.5242 R^2=0.9986
Response
Conc. (mM)

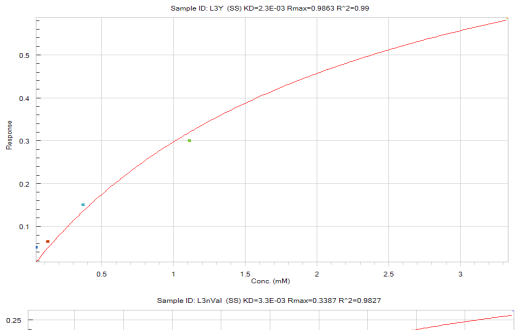
Figure AI.7 – Biolayer Interferometry data of 8-mer N-terminal Cul2 peptides with leucine replacement *versus* VBCH tested on a concentration-dependent assay.

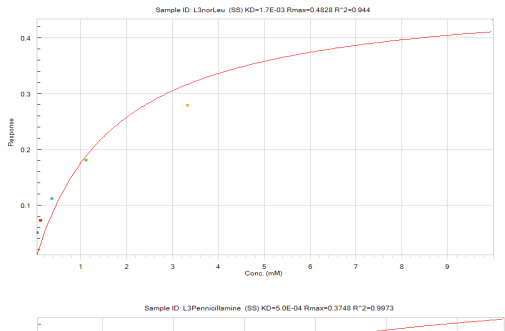


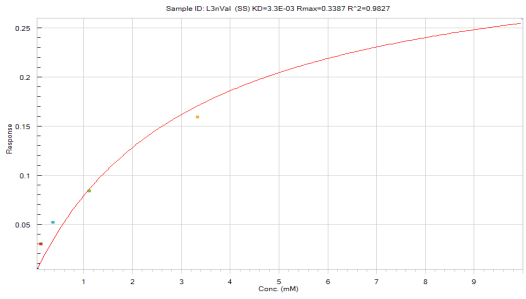

Sample ID: L3Chg (SS) KD=1.7E-02 Rmax=0.5095 R^2=0.9808
Response
Conc. (mM)

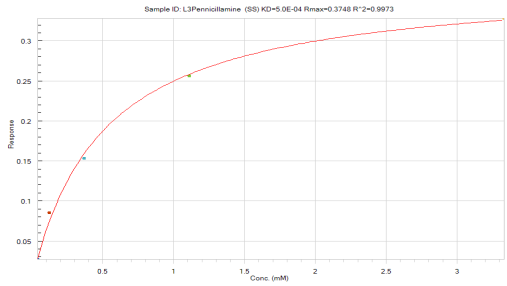

Sample ID: L3F (SS) KD=3.1E-03 Rmax=0.8554 R^2=0.9993
Response
Conc. (mM)

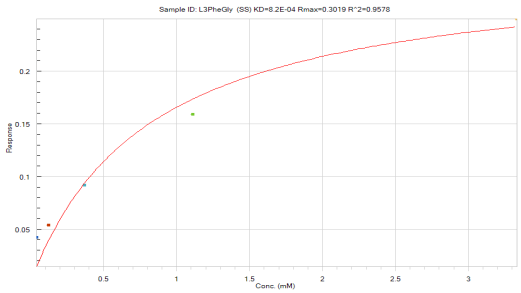

Sample ID: L3Iv (SS) KD=7.8E-03 Rmax=1.7263 R^2=0.9907
Response
Conc. (mM)

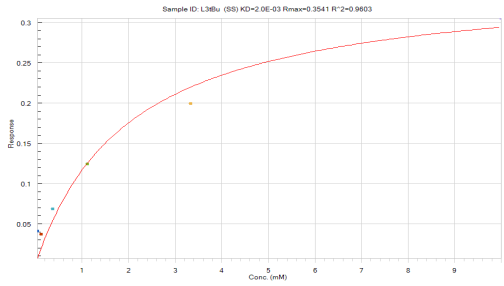

Sample ID: L3Y (SS) KD=2.3E-03 Rmax=0.9863 R^2=0.99
Response
Conc. (mM)


Sample ID: L3norLeu (SS) KD=1.7E-03 Rmax=0.4828 R^2=0.944
Response
Conc. (mM)


Sample ID: L3nVal (SS) KD=3.3E-03 Rmax=0.3387 R^2=0.9827
Response
Conc. (mM)


Sample ID: L3Penicillamine (SS) KD=5.0E-04 Rmax=0.3748 R^2=0.9973
Response
Conc. (mM)


Sample ID: L3PheGly (SS) KD=8.2E-04 Rmax=0.3019 R^2=0.9578
Response
Conc. (mM)


Sample ID: L3tBu (SS) KD=2.0E-03 Rmax=0.3541 R^2=0.9803
Response
Conc. (mM)

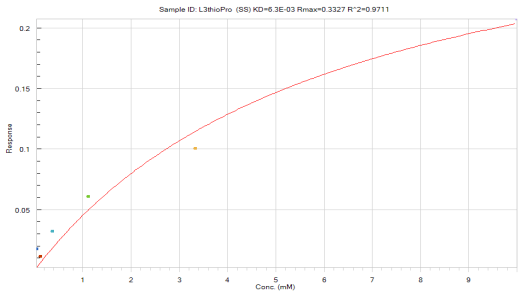
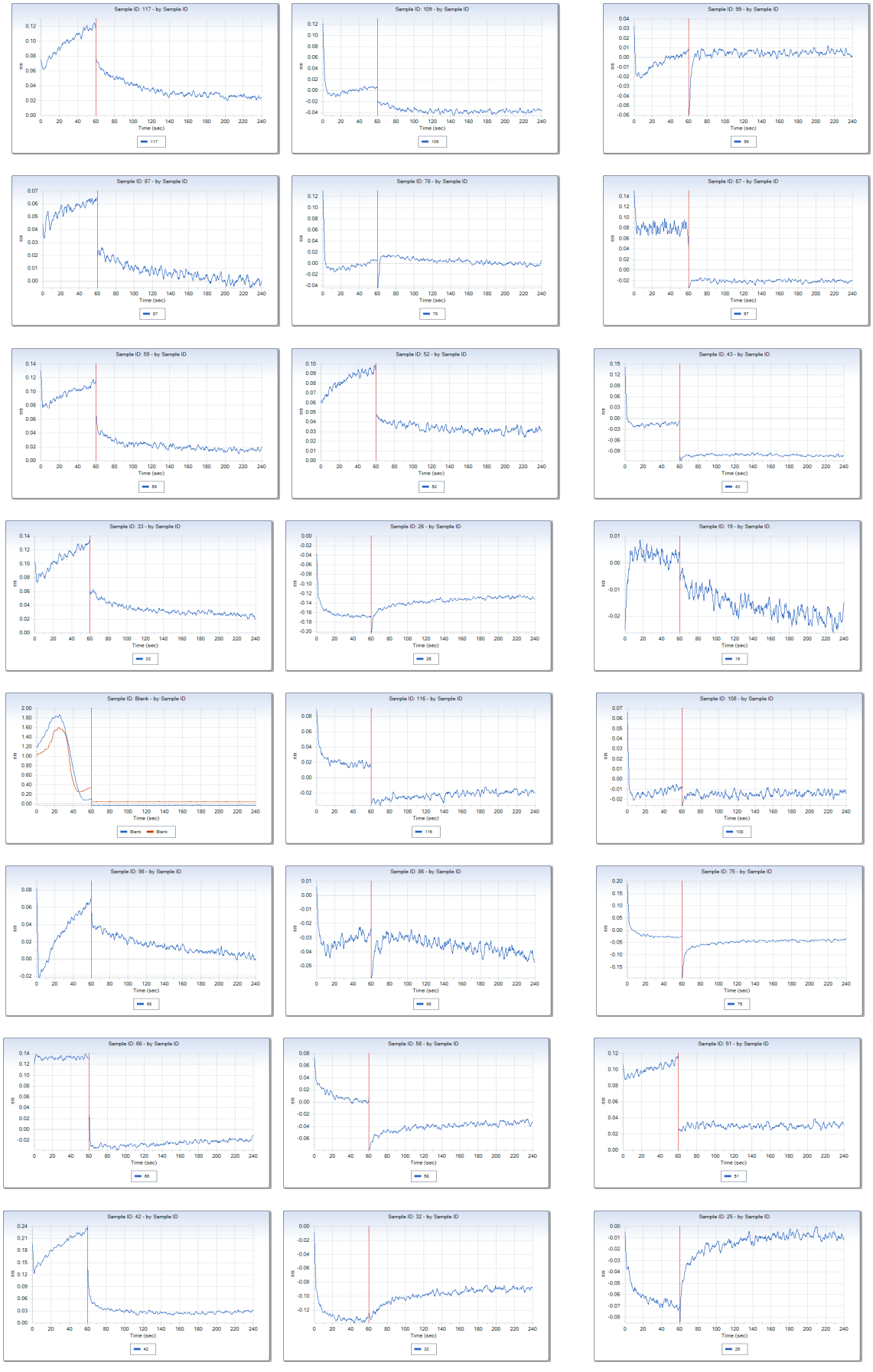
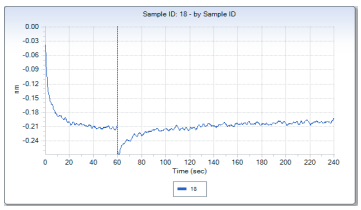
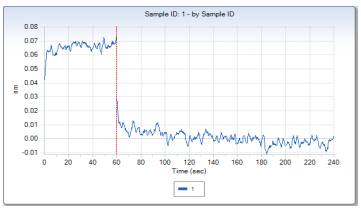
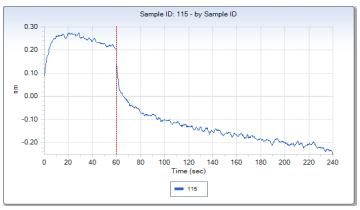

Sample ID: L3thioPro (SS) KD=6.3E-03 Rmax=0.3327 R^2=0.9711
Response
Conc. (mM)

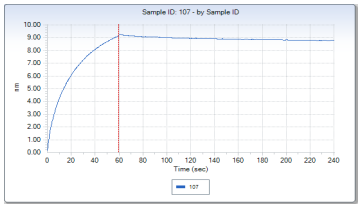
Figure AI.8 – Biolayer Interferometry data of bicyclic peptides *versus* VBC tested on a single point concentration assay.

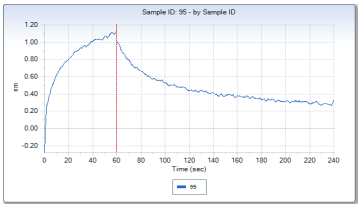


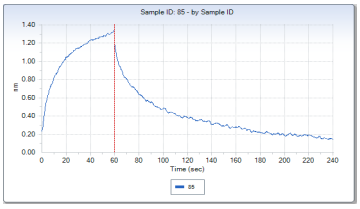

Sample ID: 18 - by Sample ID
nm
Time (sec)
18

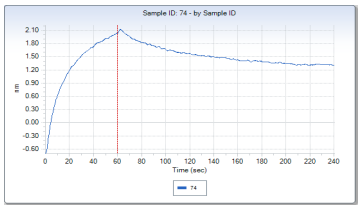

Sample ID: 1 - by Sample ID
nm
Time (sec)
1

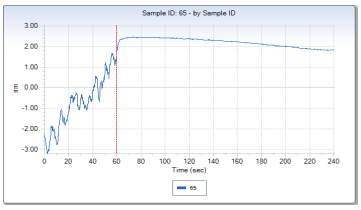

Sample ID: 115 - by Sample ID
nm
Time (sec)
115

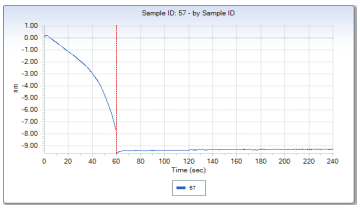

Sample ID: 107 - by Sample ID
nm
Time (sec)
107

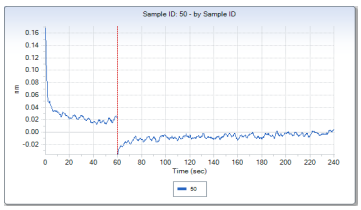

Sample ID: 95 - by Sample ID
nm
Time (sec)
95

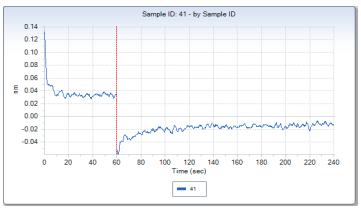

Sample ID: 85 - by Sample ID
nm
Time (sec)
85

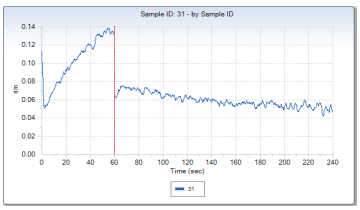

Sample ID: 74 - by Sample ID
nm
Time (sec)
74

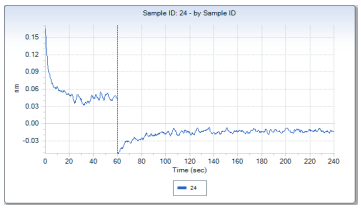

Sample ID: 65 - by Sample ID
nm
Time (sec)
65

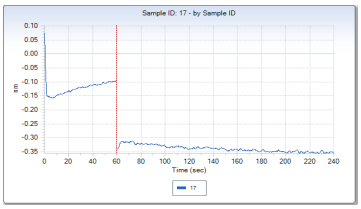

Sample ID: 57 - by Sample ID
nm
Time (sec)
57

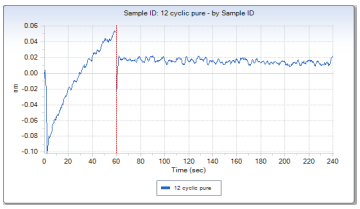

Sample ID: 50 - by Sample ID
nm
Time (sec)
50

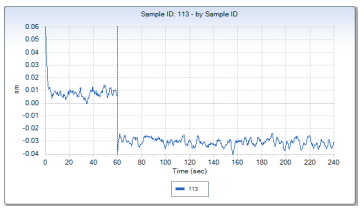

Sample ID: 41 - by Sample ID
nm
Time (sec)
41

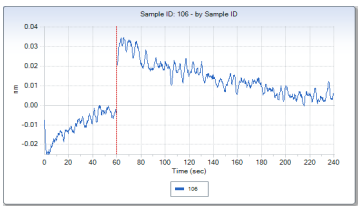

Sample ID: 31 - by Sample ID
nm
Time (sec)
31

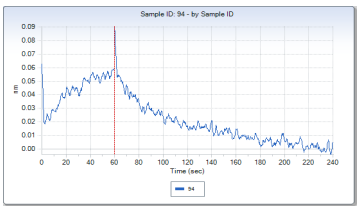

Sample ID: 24 - by Sample ID
nm
Time (sec)
24

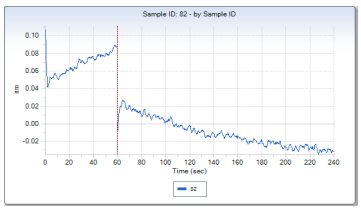

Sample ID: 17 - by Sample ID
nm
Time (sec)
17

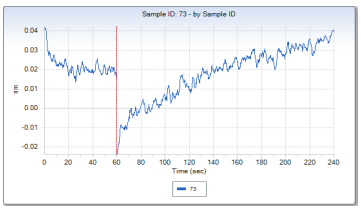

Sample ID: 12 cyclic pure - by Sample ID
nm
Time (sec)
12 cyclic pure

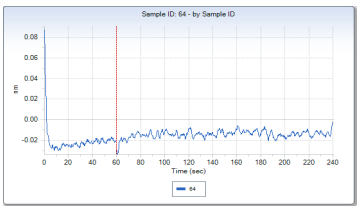

Sample ID: 113 - by Sample ID
nm
Time (sec)
113

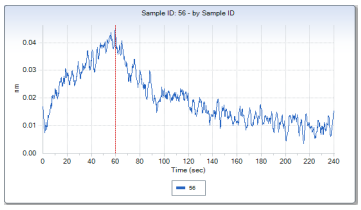

Sample ID: 106 - by Sample ID
nm
Time (sec)
106

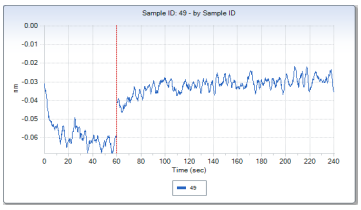

Sample ID: 94 - by Sample ID
nm
Time (sec)
94

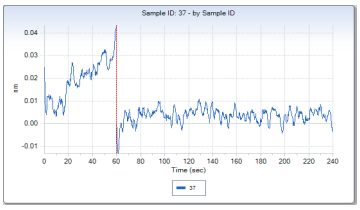

Sample ID: 82 - by Sample ID
nm
Time (sec)
82

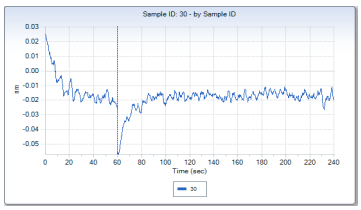

Sample ID: 73 - by Sample ID
nm
Time (sec)
73

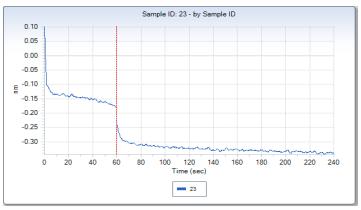

Sample ID: 64 - by Sample ID
nm
Time (sec)
64

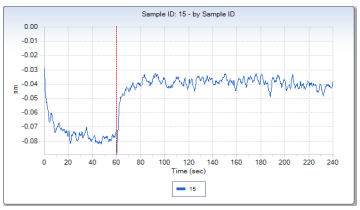

Sample ID: 56 - by Sample ID
nm
Time (sec)
56

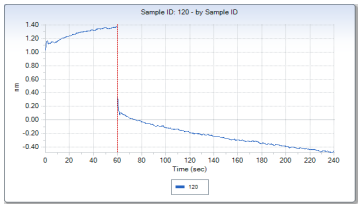

Sample ID: 49 - by Sample ID
nm
Time (sec)
49

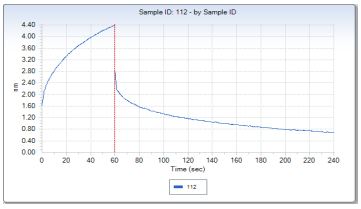

Sample ID: 37 - by Sample ID
nm
Time (sec)
37

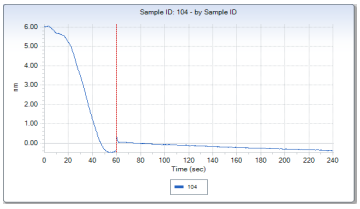

Sample ID: 30 - by Sample ID
nm
Time (sec)
30

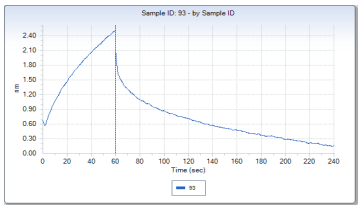

Sample ID: 23 - by Sample ID
nm
Time (sec)
23

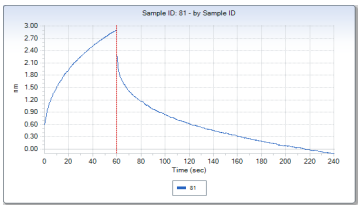

Sample ID: 15 - by Sample ID
nm
Time (sec)
15

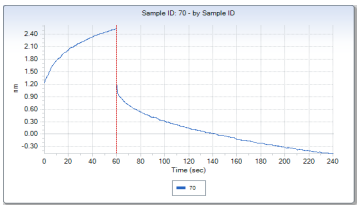

Sample ID: 120 - by Sample ID
nm
Time (sec)
120

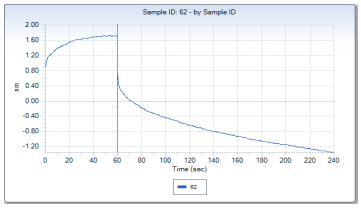

Sample ID: 112 - by Sample ID
nm
Time (sec)
112

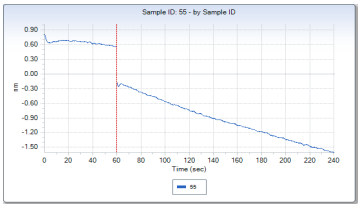

Sample ID: 104 - by Sample ID
nm
Time (sec)
104

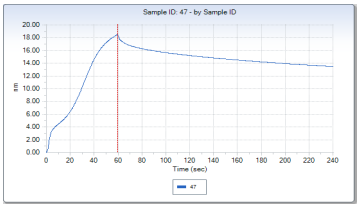

Sample ID: 93 - by Sample ID
nm
Time (sec)
93

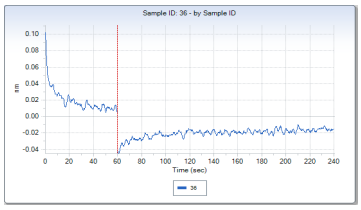

Sample ID: 81 - by Sample ID
nm
Time (sec)
81

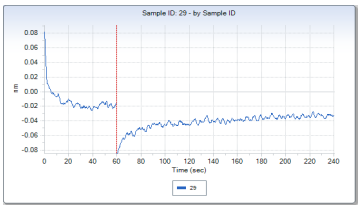

Sample ID: 70 - by Sample ID
nm
Time (sec)
70

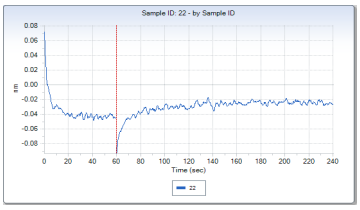

Sample ID: 62 - by Sample ID
nm
Time (sec)
62

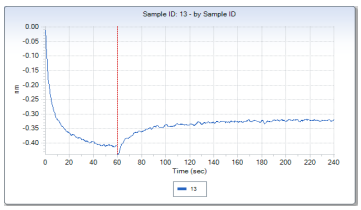

Sample ID: 55 - by Sample ID
nm
Time (sec)
55

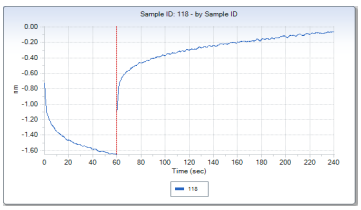

Sample ID: 47 - by Sample ID
nm
Time (sec)
47

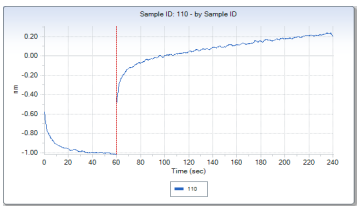

Sample ID: 36 - by Sample ID
nm
Time (sec)
36

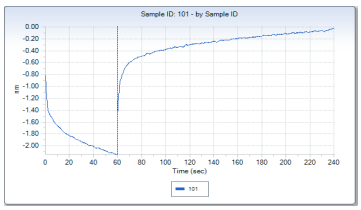

Sample ID: 29 - by Sample ID
nm
Time (sec)
29

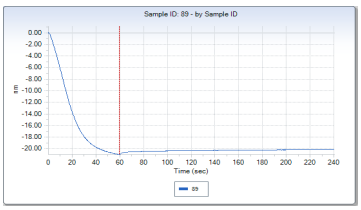

Sample ID: 22 - by Sample ID
nm
Time (sec)
22

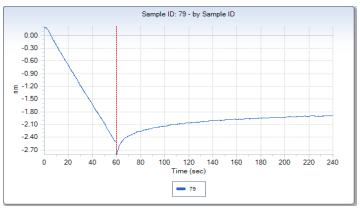

Sample ID: 13 - by Sample ID
nm
Time (sec)
13

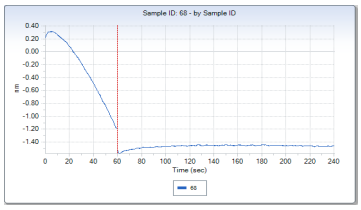

Sample ID: 118 - by Sample ID
nm
Time (sec)
118

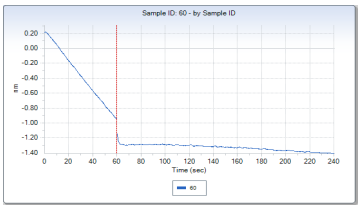

Sample ID: 110 - by Sample ID
nm
Time (sec)
110

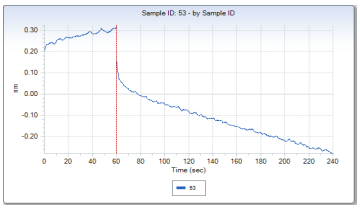

Sample ID: 101 - by Sample ID
nm
Time (sec)
101

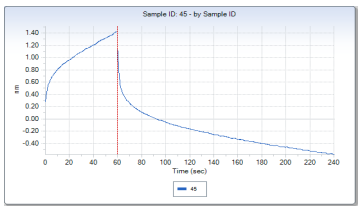

Sample ID: 89 - by Sample ID
nm
Time (sec)
89

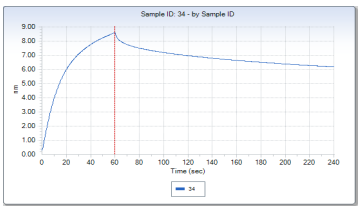

Sample ID: 79 - by Sample ID
nm
Time (sec)
79

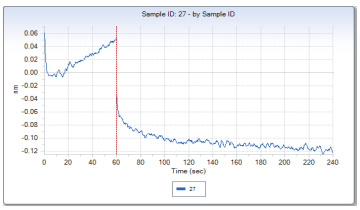

Sample ID: 68 - by Sample ID
nm
Time (sec)
68


Sample ID: 60 - by Sample ID
nm
Time (sec)
60


Sample ID: 53 - by Sample ID
nm
Time (sec)
53


Sample ID: 45 - by Sample ID
nm
Time (sec)
45


Sample ID: 34 - by Sample ID
nm
Time (sec)
34


Sample ID: 27 - by Sample ID
nm
Time (sec)
27

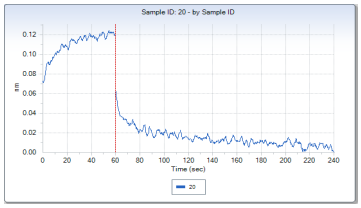
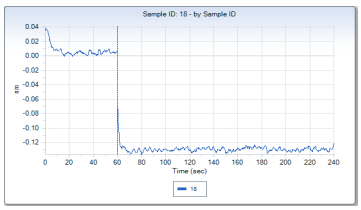
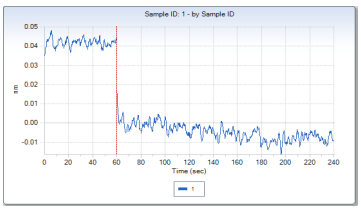
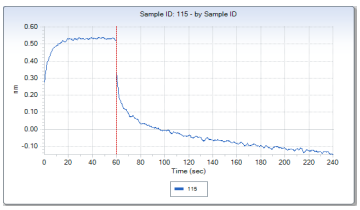
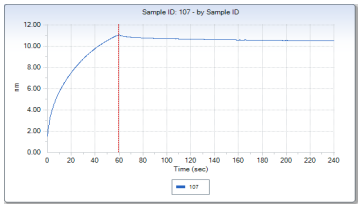

Sample ID: 20 - by Sample ID
nm
Time (sec)
20

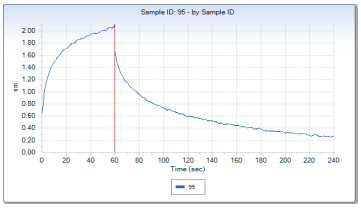
Figure AI.9 – Biolayer Interferometry data of bicyclic peptides *versus* VBCH tested on a single point concentration assay.

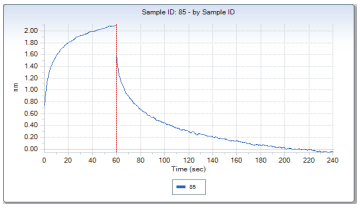

Sample ID: 18 - by Sample ID
nm
Time (sec)
18

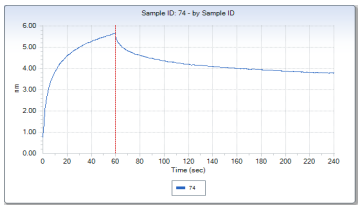

Sample ID: 1 - by Sample ID
nm
Time (sec)
1

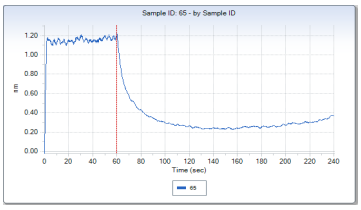

Sample ID: 115 - by Sample ID
nm
Time (sec)
115

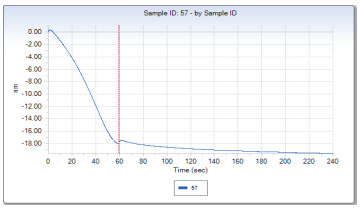

Sample ID: 107 - by Sample ID
nm
Time (sec)
107

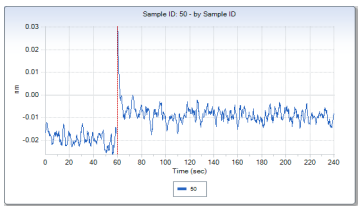

Sample ID: 95 - by Sample ID
nm
Time (sec)
95

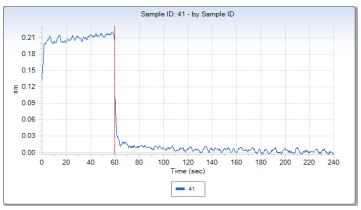

Sample ID: 85 - by Sample ID
nm
Time (sec)
85

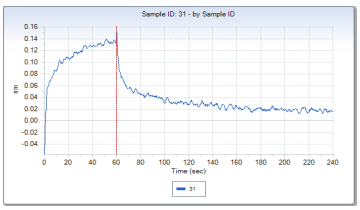

Sample ID: 74 - by Sample ID
nm
Time (sec)
74

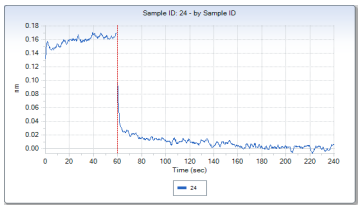

Sample ID: 65 - by Sample ID
nm
Time (sec)
65

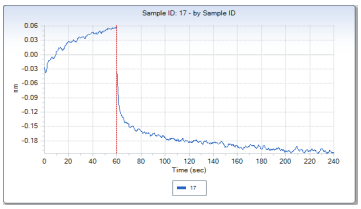

Sample ID: 57 - by Sample ID
nm
Time (sec)
57

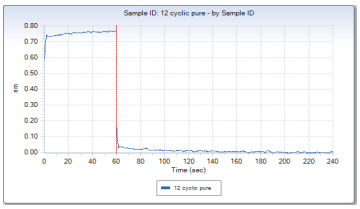

Sample ID: 50 - by Sample ID
nm
Time (sec)
50

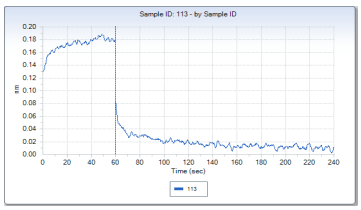

Sample ID: 41 - by Sample ID
nm
Time (sec)
41

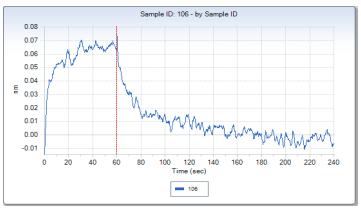

Sample ID: 31 - by Sample ID
nm
Time (sec)
31

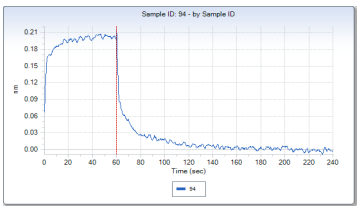

Sample ID: 24 - by Sample ID
nm
Time (sec)
24

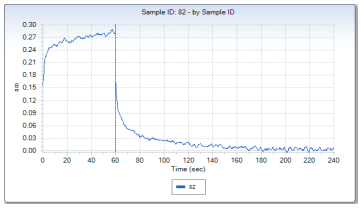

Sample ID: 17 - by Sample ID
nm
Time (sec)
17

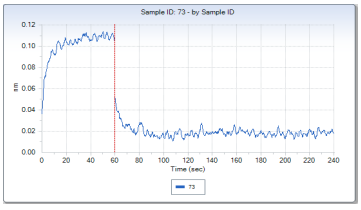

Sample ID: 12 cyclic pure - by Sample ID
nm
Time (sec)
12 cyclic pure

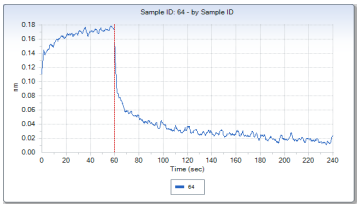

Sample ID: 113 - by Sample ID
nm
Time (sec)
113

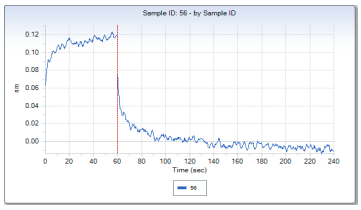

Sample ID: 106 - by Sample ID
nm
Time (sec)
106

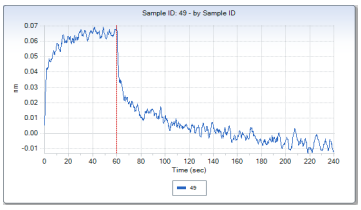

Sample ID: 94 - by Sample ID
nm
Time (sec)
94

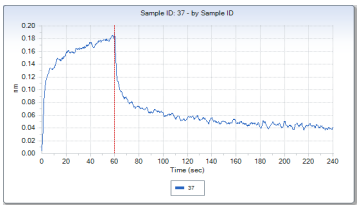

Sample ID: 82 - by Sample ID
nm
Time (sec)
82

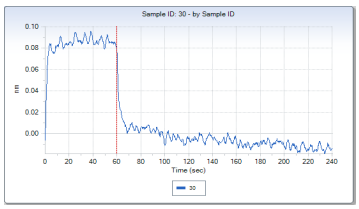

Sample ID: 73 - by Sample ID
nm
Time (sec)
73

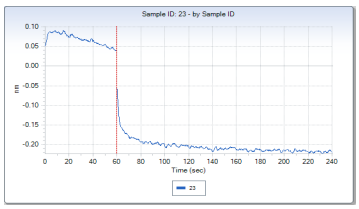

Sample ID: 64 - by Sample ID
nm
Time (sec)
64

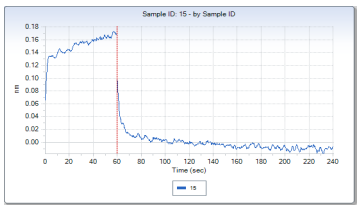

Sample ID: 56 - by Sample ID
nm
Time (sec)
56

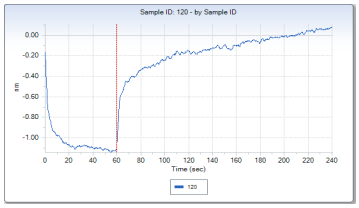

Sample ID: 49 - by Sample ID
nm
Time (sec)
49

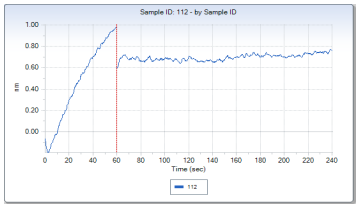

Sample ID: 37 - by Sample ID
nm
Time (sec)
37

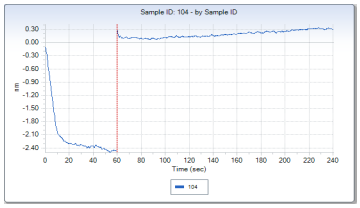

Sample ID: 30 - by Sample ID
nm
Time (sec)
30

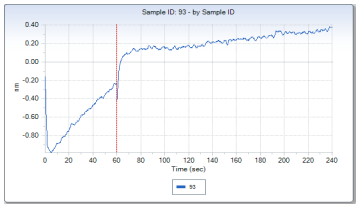

Sample ID: 23 - by Sample ID
nm
Time (sec)
23

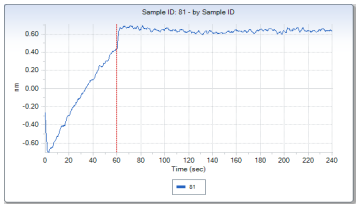

Sample ID: 15 - by Sample ID
nm
Time (sec)
15

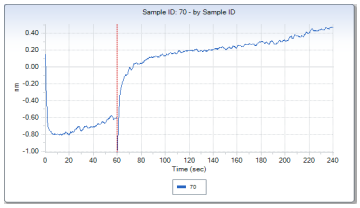

Sample ID: 120 - by Sample ID
nm
Time (sec)
120

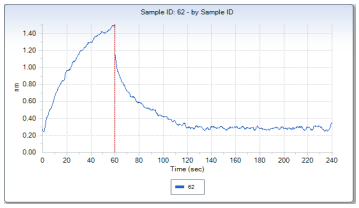

Sample ID: 112 - by Sample ID
nm
Time (sec)
112

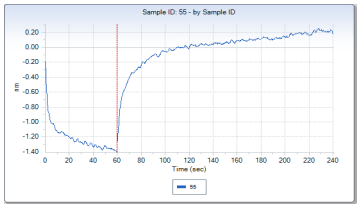

Sample ID: 104 - by Sample ID
nm
Time (sec)
104

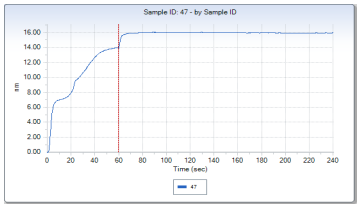

Sample ID: 93 - by Sample ID
nm
Time (sec)
93

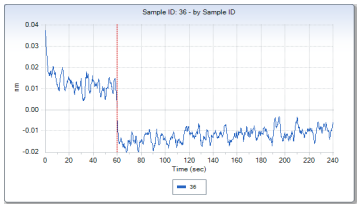

Sample ID: 81 - by Sample ID
nm
Time (sec)
81

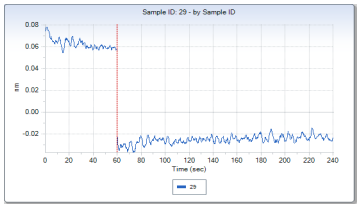

Sample ID: 70 - by Sample ID
nm
Time (sec)
70

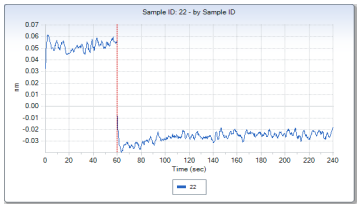

Sample ID: 62 - by Sample ID
nm
Time (sec)
62

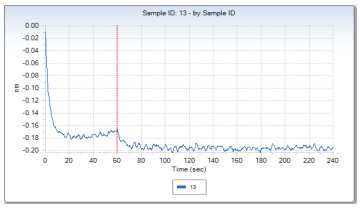

Sample ID: 55 - by Sample ID
nm
Time (sec)
55

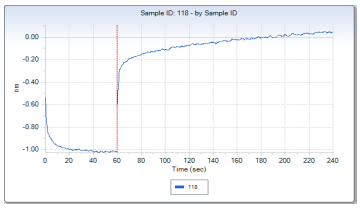

Sample ID: 47 - by Sample ID
nm
Time (sec)
47

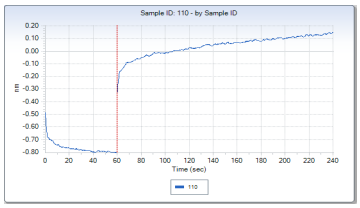

Sample ID: 36 - by Sample ID
nm
Time (sec)
36

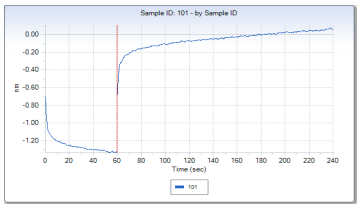

Sample ID: 29 - by Sample ID
nm
Time (sec)
29

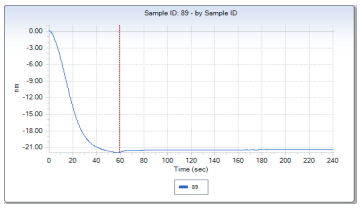

Sample ID: 22 - by Sample ID
nm
Time (sec)
22

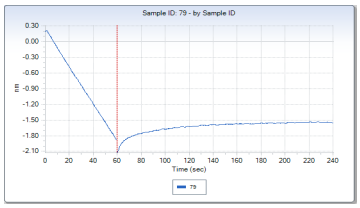

Sample ID: 13 - by Sample ID
nm
Time (sec)
13

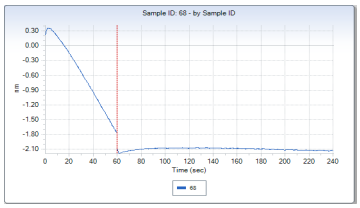

Sample ID: 118 - by Sample ID
nm
Time (sec)
118

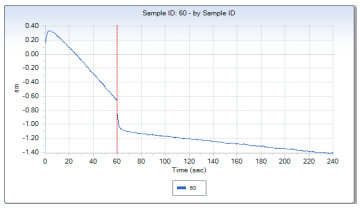

Sample ID: 110 - by Sample ID
nm
Time (sec)
110

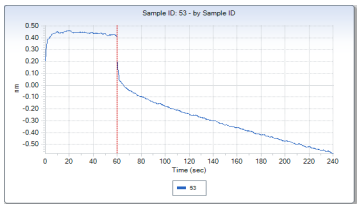

Sample ID: 101 - by Sample ID
nm
Time (sec)
101

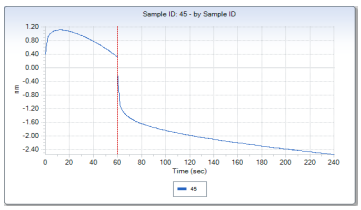

Sample ID: 89 - by Sample ID
nm
Time (sec)
89

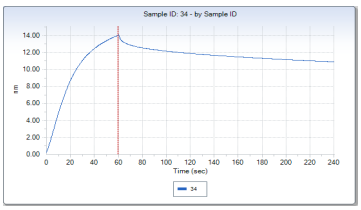

Sample ID: 79 - by Sample ID
nm
Time (sec)
79

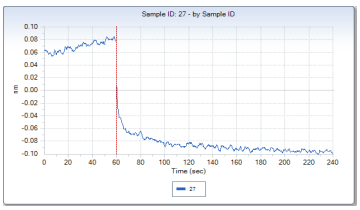

Sample ID: 68 - by Sample ID
nm
Time (sec)
68


Sample ID: 60 - by Sample ID
nm
Time (sec)
60


Sample ID: 53 - by Sample ID
nm
Time (sec)
53


Sample ID: 45 - by Sample ID
nm
Time (sec)
45


Sample ID: 34 - by Sample ID
nm
Time (sec)
34


Sample ID: 27 - by Sample ID
nm
Time (sec)
27

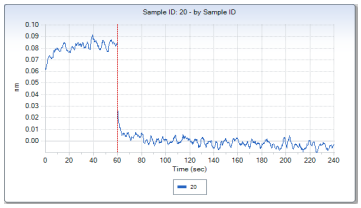
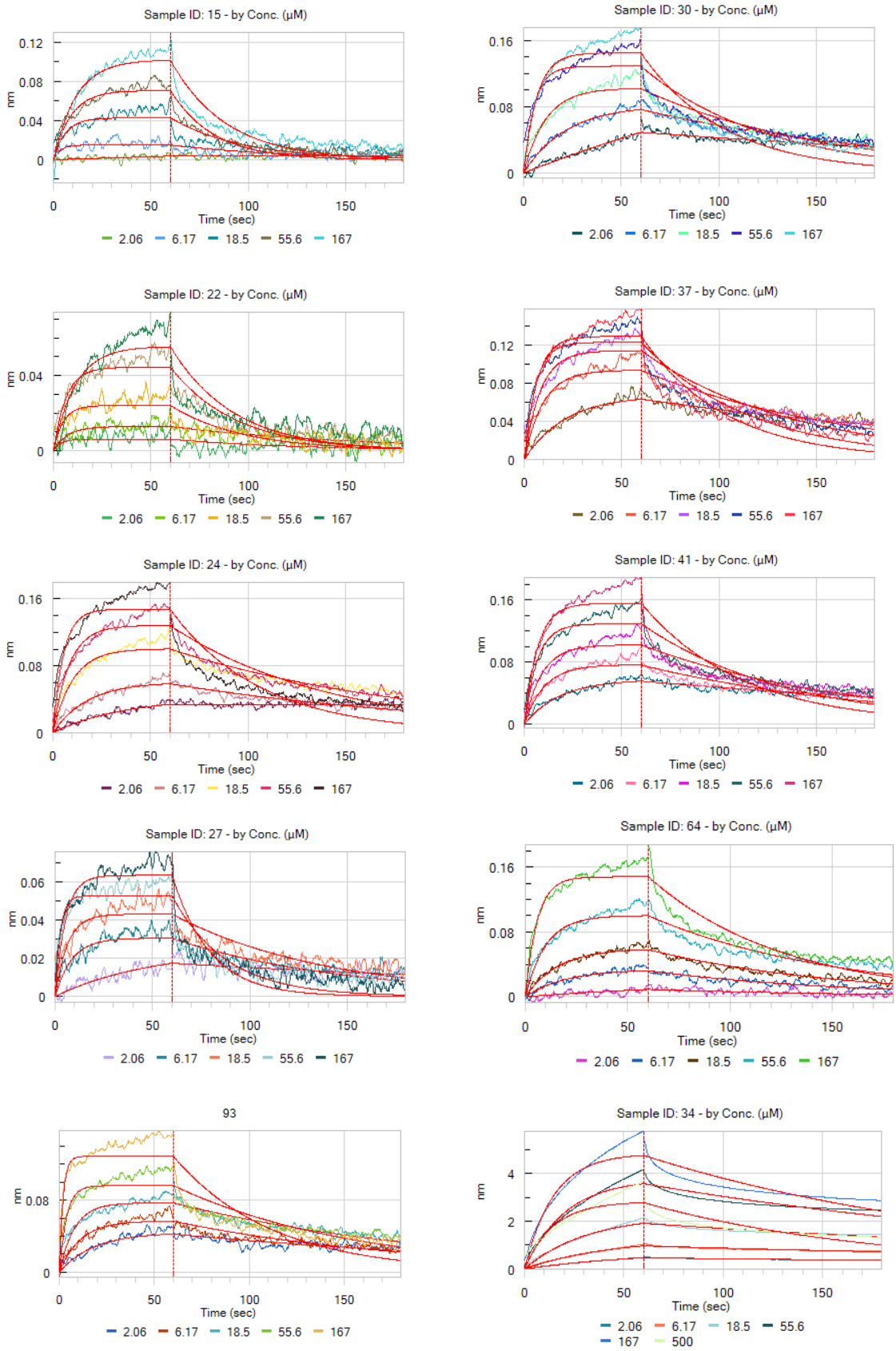
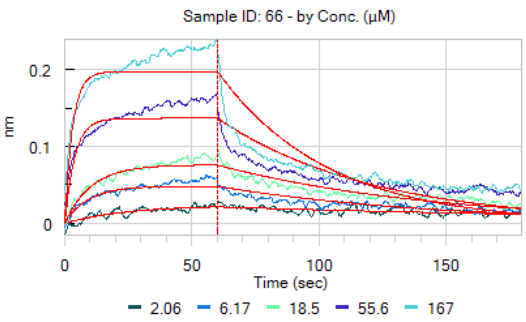
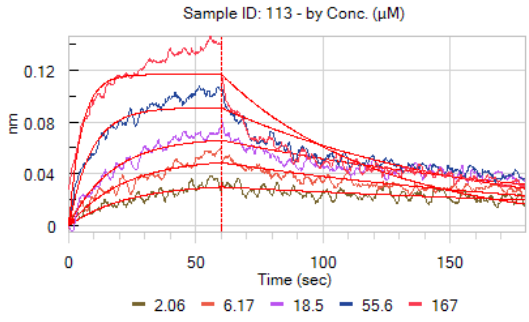
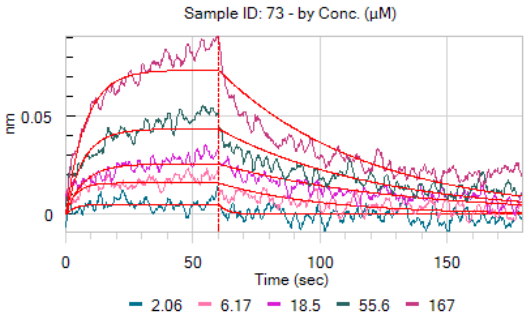

Sample ID: 20 - by Sample ID
nm
Time (sec)
20

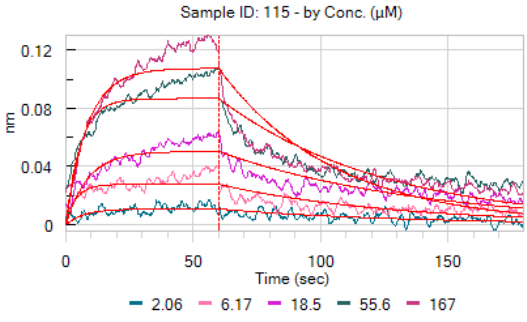
Figure AI.10 – Biolayer Interferometry data of bicyclic peptides *versus* VBC tested on a concentration-dependent assay.

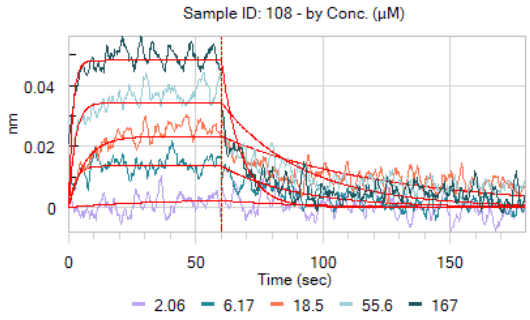


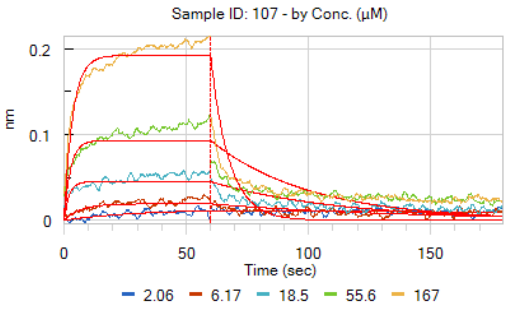

Sample ID: 66 - by Conc. (µM)
nm
0.2
0.1
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167

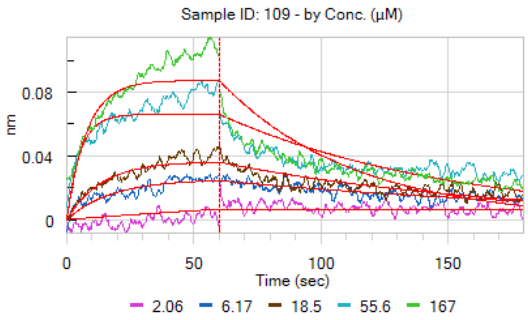

Sample ID: 113 - by Conc. (µM)
nm
0.12
0.08
0.04
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167

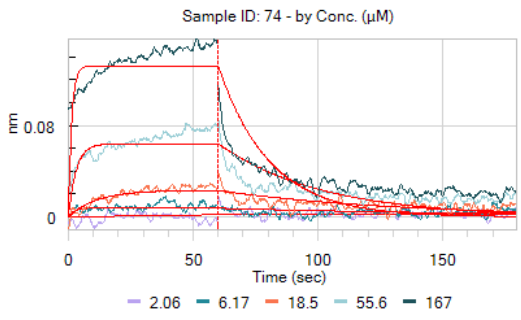

Sample ID: 73 - by Conc. (µM)
nm
0.05
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167


Sample ID: 115 - by Conc. (µM)
nm
0.12
0.08
0.04
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167


Sample ID: 108 - by Conc. (µM)
nm
0.04
0.02
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167


Sample ID: 107 - by Conc. (µM)
nm
0.2
0.1
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167


Sample ID: 109 - by Conc. (µM)
nm
0.08
0.04
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167


Sample ID: 74 - by Conc. (µM)
nm
0.08
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167

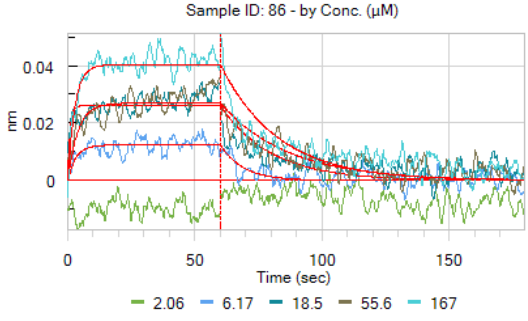
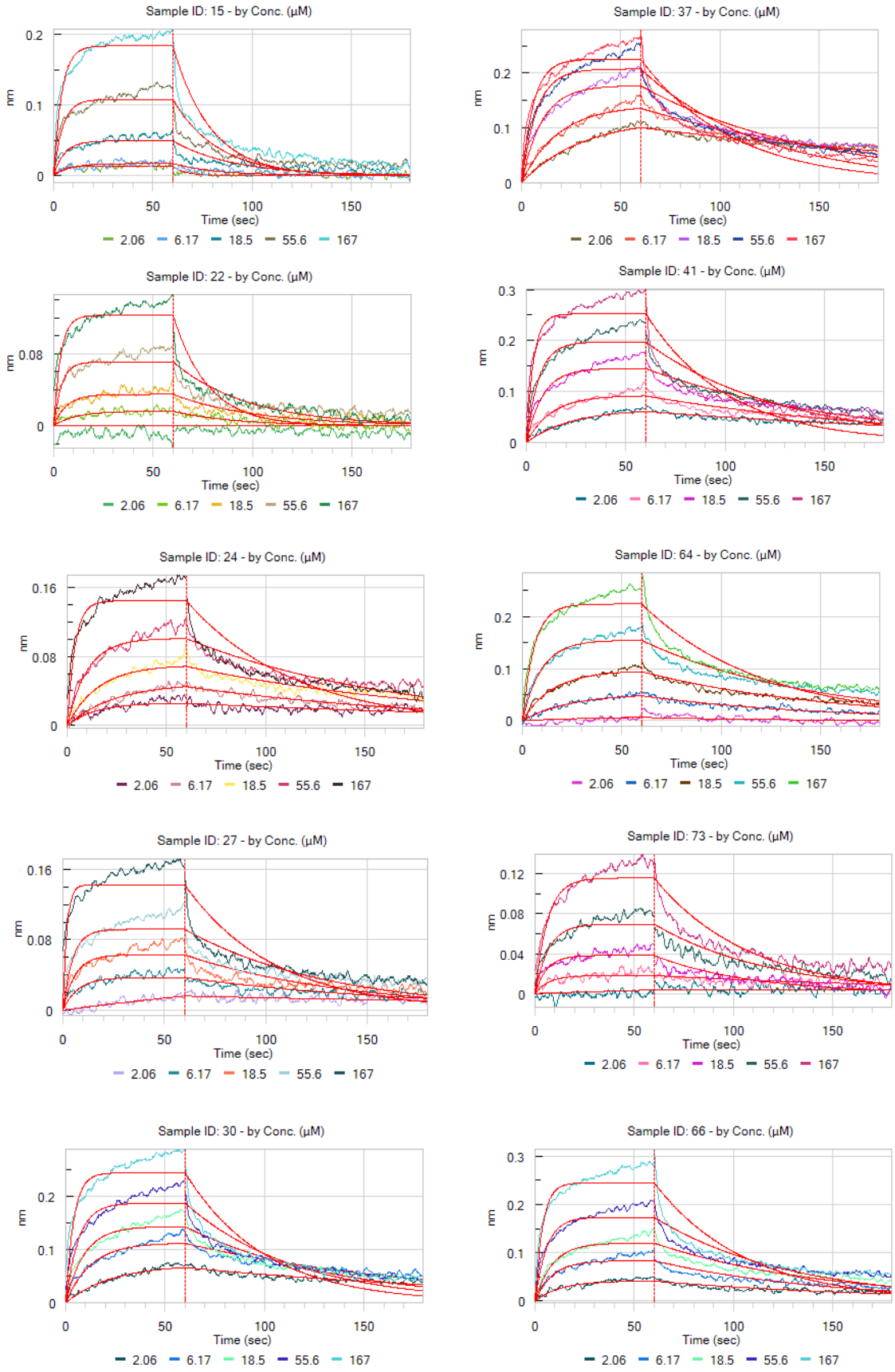
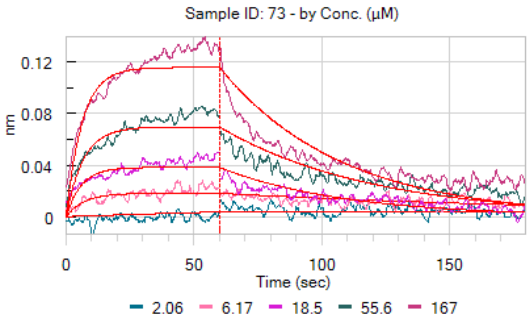
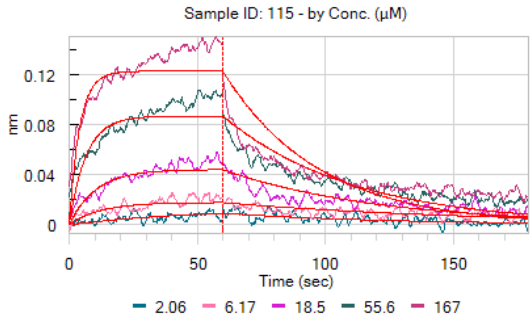
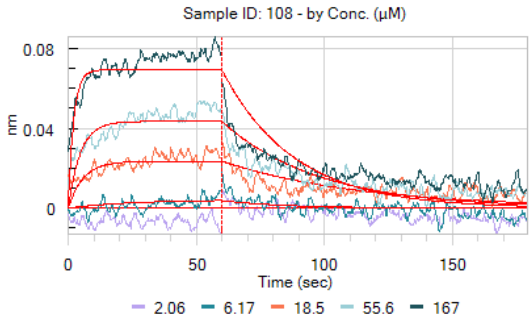

Sample ID: 86 - by Conc. (µM)
nm
0.04
0.02
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167

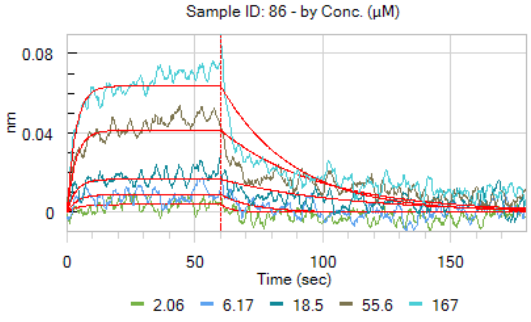
Figure AI.11 – Biolayer Interferometry data of bicyclic peptides *versus* VBCH tested on a concentration-dependent assay.

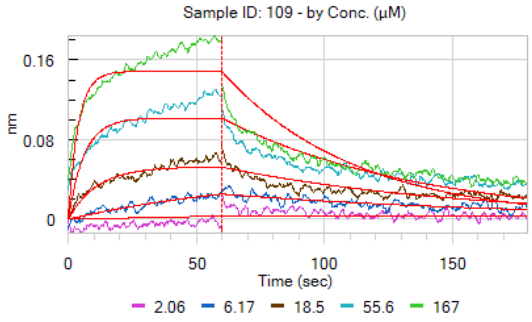


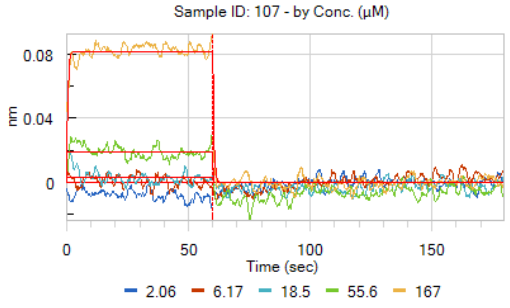

Sample ID: 73 - by Conc. (µM)
nm
0.12
0.08
0.04
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167

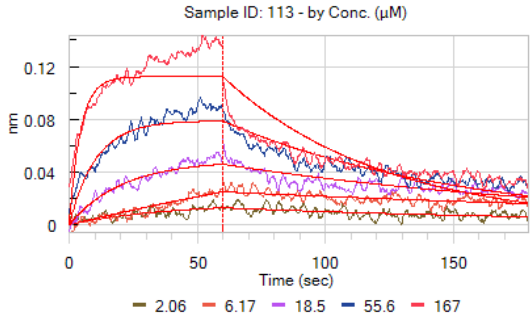

Sample ID: 115 - by Conc. (µM)
nm
0.12
0.08
0.04
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167

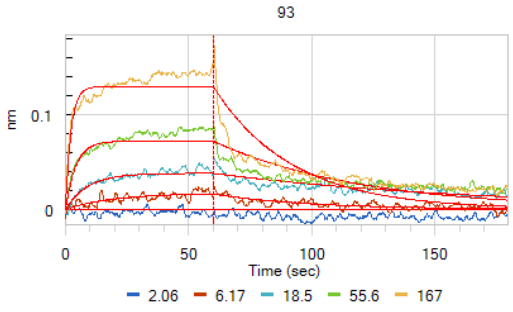

Sample ID: 108 - by Conc. (µM)
nm
0.08
0.04
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167

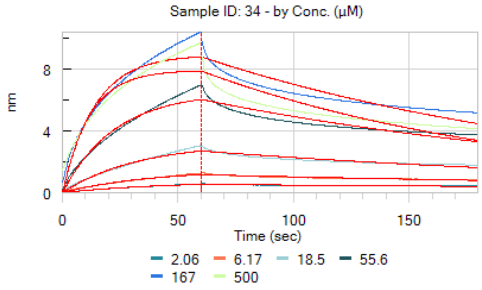

Sample ID: 86 - by Conc. (µM)
nm
0.08
0.04
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167


Sample ID: 109 - by Conc. (µM)
nm
0.16
0.08
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167


Sample ID: 107 - by Conc. (µM)
nm
0.08
0.04
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167


Sample ID: 113 - by Conc. (µM)
nm
0.12
0.08
0.04
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167


93
nm
0.1
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167


Sample ID: 34 - by Conc. (µM)
nm
8
4
0
0
50
100
150
Time (sec)
2.06
6.17
18.5
55.6
167
500

APPENDIX II